MEDICAL ROOM
MEDICAL LIBRARY

The JOURNAL *of* EXPERIMENTAL EDUCATION

Volume I | SEPTEMBER, 1932 | Number 1

CONTENTS

The Existence, Measurement, and Significance of a Speed Factor in the Abilities of Public School Children: W. LINE 1

An Experimental Study of Functional Learning: A. S. BARR *and* JOHN S. PARK 9

The Relative Contribution of Certain Factors to Individual Differences in Arithmetical Problem Solving Ability: M. D. ENGLEHART 19

An Experimental Analysis of Reading Performance: OSCAR F. LITTERER .. 28

A Study of the Relative Amounts of Transfer Resulting from Three Methods of Study: JOHN B. CLULEY 34

Comparison of Two Methods of Studying History: JOHN J. DYNES 42

$5.00 A YEAR | PUBLISHED QUARTERLY | $1.50 A COPY

EDWARDS BROTHERS, INC.

LITHOPRINTERS AND PUBLISHERS — ANN ARBOR, MICHIGAN

Application filed for entry as second class matter.

JOURNAL OF EXPERIMENTAL EDUCATION

EDITORIAL BOARD

A. S. Barr, Chairman, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Carter V. Good, Professor of Education, University of Cincinnati, Cincinnati, Ohio. Editorially responsible for materials on supervision and psychology of learning and teaching.

Henry Harap, Professor of Education, Western Reserve University, Cleveland, Ohio. Editorially responsible for materials on experimental studies of curriculum construction.

Walter S. Monroe, Professor of Education, University of Illinois, Urbana, Illinois. Editorially responsible for materials on measurements, statistics, and methods of experimental research.

George D. Stoddard, Director, Child Welfare Research Station, State University of Iowa, Iowa City, Iowa. Editorially responsible for materials on child welfare, guidance, and development.

CONTRIBUTING EDITORS

Harry J. Baker, Director Psychological Clinic, Detroit Public Schools, Detroit, Michigan.

William A. Brownell, Professor of Experimental Education, Duke University, Durham, North Carolina.

W. E. Blatz, Nursery School Division, University of Toronto, Toronto 5, Canada.

Leo J. Brueckner, Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Herbert B. Bruner, Professor of Education, Teachers College, Columbia University, New York City.

Barbara S. Burks, Department of Psychology, Stanford, University, California.

Ellsworth Collings, Dean, College of Education, University of Oklahoma, Norman, Oklahoma.

Philip W. L. Cox, Professor of Secondary Education, New York University, New York City.

Edgar A. Doll, Director, Department of Research, Training School, Vineland, New Jersey.

Harl R. Douglass, Professor of Secondary Education, University of Minnesota, Minneapolis, Minnesota.

Jack W. Dunlap, Professor of Education, Fordham University, 233 Broadway, New York City.

Paul H. Furfey, Professor of Psychology, Catholic University of America, Washington, D. C.

Florence L. Goodenough, Professor, Institute of Child Welfare, University of Minnesota, Minneapolis, Minnesota.

Karl J. Holzinger, Professor of Education, University of Chicago, Chicago, Illinois.

L. Thomas Hopkins, Curriculum Specialist, Lincoln School of Teachers College, 425 West 123rd Street, New York City.

C. L. Huffaker, Professor of Education, University of Oregon, Eugene, Oregon.

Kai Jensen, Assistant Professor of Psychology, University of Wisconsin, Madison, Wisconsin.

Harold E. Jones, Director of Research, Institute of Child Welfare, University of California, Berkeley, California.

Edward A. Lincoln, Assistant Professor of Education, Graduate School of Education, Harvard University, Cambridge, Massachusetts.

E. F. Lindquist, Professor of Education, State University of Iowa, Iowa City, Iowa.

Lois Hayden Meek, Director, Child Development Institute, Teachers College, Columbia University, New York City.

C. W. Odell, Associate Professor of Education, University of Illinois, Urbana, Illinois.

Willard C. Olson, Associate Professor of Education, University of Michigan, Ann Arbor, Michigan.

W. E. Peik, Associate Professor of Education, University of Minnesota, Minneapolis, Minnesota.

S. L. Pressey, Professor of Educational Psychology, Ohio State University, Columbus, Ohio.

W. H. Pyle, Professor of Educational Psychology, Detroit Teachers College, Detroit, Michigan.

Clarence E. Ragsdale, Assistant Professor of Education, University of Wisconsin, Madison, Wisconsin.

H. H. Remmers, Professor of Education and Psychology, Purdue University, Lafayette, Indiana.

G. M. Ruch, Professor of Education, University of California, Berkeley, California.

Earl U. Rugg, Head, Department of Education, Colorado State Teachers College, Greeley, Colorado.

Peter Sandiford, Professor of Educational Psychology, Ontario College of Education, University of Toronto, Toronto, Canada.

Douglas E. Scates, Director, School Research, Cincinnati Public Schools, Cincinnati, Ohio.

Raleigh Schorling, Professor of Education, University of Michigan, Ann Arbor, Michigan.

Mandel Sherman, Professor of Education, University of Chicago, Chicago, Illinois.

Helen Thompson, Research Associate, Yale University, New Haven, Connecticut.

Herbert A. Toops, Research Associate, Ohio State University, Columbus, Ohio.

T. L. Torgerson, Assistant Professor of Education, University of Wisconsin, Madison, Wisconsin.

M. R. Trabue, Executive Secretary of Committee on Individual Diagnosis and Training, Institute for Employment Stabilization Research, University of Minnesota, Minneapolis, Minnesota.

R. W. Tyler, Professor of Education, Ohio State University, Columbus, Ohio.

Douglas Waples, Graduate Library School, University of Chicago, Chicago, Illinois.

Beth L. Wellman, Research Associate Professor, Child Welfare Research Station, University of Iowa, Iowa City, Iowa.

PURPOSE

The *Journal of Experimental Education* is a technical journal for technical workers interested in the experimental study of education. The materials published are limited to: (1) reports of experimental studies of the child, the curriculum, and methods of instruction; and (2) critical discussions and researches relating to the measurements, statistics, and methods of experimental research.

A PROJECT IN COOPERATIVE PRINTING AND SELLING

The *Journal of Experimental Education* is a project in cooperative printing and selling charging a contributor's fee of $2.50 per page of 800 words calculated on the basis of a fixed annual budget and dividing all income beyond this amount pro rata according to the number of pages of material printed per contributor. The budget for 1932-33 is $2,000 or the equivalent of four hundred subscriptions at $5.00 each. If this number of subscriptions is secured the contributor's fee is returned; if there are fewer than four hundred subscriptions, a fractional portion of the contributor's fee is returned after the publication costs have been covered; if there are additional subscriptions the income is divided pro rata according to the number of pages printed per contributor.

A JOURNAL AND A YEARBOOK

The *Journal of Experimental Education* is published quarterly in September, December, March, and June, each issue being devoted to a special phase of education: the September issues to supervision and the psychology of learning and teaching; the December issues to child welfare, guidance and development; the March issues to measurements, statistics, and the methods of experimental research; and the June issues to curriculum construction. The Journal is both a journal and a series of yearbooks.

SUBSCRIPTIONS

The subscription price is $5.00 per year, in advance; Canada, $6.00 a year; other foreign countries, $7.00 a year. The price for single copies will vary from issue to issue, depending upon the size of the issue. The price for single copies is given on the front cover of each issue. Lots of 25 copies or more of single issues may be purchased at a ten per cent discount. Each issue of the Journal will contain approximately 100 pages of lithoprinting or a total of 400 pages of lithoprinting per year. Four hundred pages of lithoprinting is approximately equivalent to 800 pages of printing.

EDITORIAL MATTERS

Address all communications on editorial matters to A. S. Barr, Professor of Education, University of Wisconsin, Madison, Wisconsin. Manuscripts intended for publication in any one of the above named issues of the Journal should reach the editor of that issue at least sixty days prior to the date of publication. Postage for the return of the manuscript should be enclosed.

BUSINESS MATTERS

Address all communications on business matters, subscriptions, reprints, contributor's fees, advertising matters, etc., to Edwards Brothers, Lithoprinters and Publishers, Ann Arbor, Michigan.

The JOURNAL of EXPERIMENTAL EDUCATION

Volume I | September 15, 1932 | Number 1

THE EXISTENCE, MEASUREMENT AND SIGNIFICANCE OF A SPEED FACTOR IN THE ABILITIES OF PUBLIC SCHOOL CHILDREN[1]

By W. Line and E. Kaplan
University of Toronto

I. Introduction

One of the important controversies in the field of mental measurements is concerned with whether ordinary tests of intelligence are chiefly measures of speed or of power. "Power" has been defined to some purpose; objectively, in terms of the difficulty of test items; subjectively, in relation to the complexity of the thought-processes involved. In the case of "speed," the objective definition seems fairly clear, in that it refers to the time taken to perform a given task. Difference of opinion exists, however, with regard to the psychological interpretation of the term because of its several possible aspects. The study here reported deals with an examination of intelligence test scores, with a view to seeing if some one aspect of speed comes to light. It then proceeds to investigate that aspect further by means of specially-constructed tests.

II. Historical Review

The term "speed" has been used in psychological literature with various meanings. It is not surprising to find, therefore, many conflicting statements. As an instance of this we may take the opposing views that speed is unrelated to intelligence, and that speed and intelligence are synonymous as characteristic of an individual's ability.

The first view, that speed is unrelated to intelligence, is implied in the popular assumption that an individual may be slow, yet profound, able, and accurate. It is the view that speed is either an inherent idiosyncrasy which is not of vital importance in most walks of life, where "power" is required; or it is the result of a particular environmental demand. To some extent, criticisms of tests of intelligence on the ground that they are only tests of speed, mirror the same opinion. Researches from which conclusions are drawn in support of this view have been reported by Highsmith (5) and Walters (11).

The second view identifies speed with intelligence. To be quick is to be intelligent, while slowness is synonymous with dullness. The work of Boring and Peak (2), and of Gilbert (3) may be cited as pertinent to this conception of speed. Spearman to some extent (10) similarly subscribes to the same notion. At the same time, however, he has differentiated experimentally between an "attitude to or preference for speed" and "speed of cognition." The evidence offered by him points to the interchangeability of "speed of cognition" and "power" (or "goodness") and both, he says, are manifestations of g. Bernstein's work (1) corroborates this general finding. So, too, Kelley (6) and Hargreaves (4) have succeeded in applying the techniques of factor-analysis to this field and have indicated the independence of the two aspects of speed here mentioned. Reymert's demonstration (8) of a group-factor in the field of reaction times may be likewise in harmony with the foregoing.

The results obtained in the American Army experiment, and in some related researches, might be illuminated further if interpreted along the lines suggested by this factorial approach. We refer particularly to the reports of May and Terman (7), and of Ruch and

1. Acknowledgment is here made of the generous cooperation of Mr. S. Richardson, and his staff, among whom we would mention especially Miss Raymond and Miss Sloan. Without their assistance this study would not have been possible.

Koerth (9). With this point of view before us, the first experiment was conducted.

III. The First Experiment

The National Intelligence Test (Form A) was given to a class of 44 Senior Fourth pupils at a Toronto Public School. The directions used in giving the test were altered from the standardized form only to the degree necessary to obtain scores for half-time, standard time, and extended time, the latter being one and one-half times the standard period. A practice score was also obtained by using the results of the exercise preceding every test-proper in the National Scale.

Selecting from the results only those that are of direct interest here we may note that all but three subjects increased their scores when the time was extended. None, however, obtained full marks even with extended time, so that no subjects were handicapped in this period by the fact of having already completed the test. In general, the individuals scoring low in the standard time period, gained more than those scoring high. Further analysis of the results showed that subjects obtaining high scores in the half-time-period continued to gain more during the second half of the standard time, than the low-scoring individuals. The low-scoring individuals however, gained more during the extended time period.

It appears that the duller pupils had not reached their limit to as great a degree as had the brighter by the end of the standard time; that is, when they were given more time, they improved in relation to the brighter ones. This may be due to one or more of several things:

(a) The bright students were doing harder problems during the extended time. Therefore the points obtained by bright and dull were not comparable. This is obviously true, since the test items are arranged approximately in order of difficulty.

(b) The bright students warm up or become orientated more speedily than the dull, thus giving the dull a disadvantage in short-time tests. That this is true seems indicated by the fact that the bright students vary from the dull students more on the first-half than on the second-half time period. This decrease in variability, however, may be due to the differing degree of difficulty of the items attacked by the two groups of subjects as the test period proceeds.

(c) The easier material at the beginning of test-groups shows differences not so much of "intellectual speed" as of speed of another kind,--an attitude toward, or preference for, speed. When the material gets harder, this attitude toward speed sinks into less importance compared with the difficulty-speed (thought-speed) as a differentiating factor. If this were true, a speed factor other than g might show itself in practice and half-time scores.

An examination of the correlations given in Table I suggests that this latter interpretation may be warranted:

TABLE I

INTERCORRELATIONS BETWEEN SCORES OBTAINED UNDER DIFFERENT CONDITIONS OF TESTING

	Practice	First Half	Second Half	Extended Time Minus Standard	Standard
Practice	-	.53(.07)*	.28(.09)	-.37(.09)	.51(.08)
First half	.53	-	.16(.10)	-.24(.10)	.80(.04)
Second half	.28	.16	-	-.41(.08)	.63(.06)
Extd. minus Standard	-.37	-.24	-.41	-	-.44(.08)
Standard	.51	.80	.63	-.44	-

*Note:--The Probable Errors are given in Parentheses.

With so small a group, tetrad analysis can be undertaken only with extreme caution. The tentative suggestion that the specific correlation between the abilities represented by the practice and the first-half scores is due to an attitude toward speed, or alertness, it being effective in both. That such a factor exists, will be demonstrated more fully in the next section. Here it may be added that the attitude for speed might conceivably work against success in difficult tasks, since the tendency would be to pass on to the next item as soon as hindrance is felt, and so the number of successes in the later section of the test would be smaller.

IV. The Second Experiment

In the preceding analysis of intelligence test scores obtained under different time conditions, the suggestion seemed to be warranted that, with the easier material, speed other than speed of thinking played a part in the results. This suggestion arose from the fact that a group factor appeared when two sections of easy tasks were considered. It was therefore decided to construct tests of very simple material and to give these under conditions of limited time, so that such a speed factor might conceivably be detected more surely, and examined in relation to some other abilities.

A. Procedure

A battery of tests was devised composed of the following:

1. Substitution--numbers for symbols.
2. Numbers--same and different.
3. Easy Word Analogies.
5. Simple Addition.
6. Words--same and opposite.
7. Serial Addition.

The aim in the construction of these tests was to keep the items of approximately equal difficulty throughout each test, and simple enough to be applied to pupils in the Junior Third Grade. The purpose was to obtain a measure of speed, so that "power" was practically excluded by the simplicity of the tasks. A dotting test (Test 4) was included as a means of indicating any significant change in attitude towards the experiment.

The material was gathered together and submitted to one class of Junior Third pupils in a preliminary experiment to make sure that the material was suitable and the instructions definite. Time allotments also were judged on the basis of this trial. The pupils who acted as subjects at this time were not involved in later experimentation.

Two groups of Junior Third Pupils were used as subjects. These groups were equivalent as classified on the basis of metal and chronological age, I.Q., scholastic attainments, and behaviour records. The tests were given on four consecutive days at the same hour of the morning. Group A, tested at 9:10, took the tests in order, 1, 2, 3, 4, 5, 6, and 7. With Group B, tested at 9:35, the order was reversed. As far as possible the conditions of motivation were kept constant for the various repetitions and for the two groups, by means of the definiteness of the instructions and constancy of the examiner. The experimental checks used in relation to this point will be discussed later.

The criterion of intelligence was the Pintner Group Test for Intelligence, Form A. (1932 edition). The number of subjects for which complete results were obtained was sixty-four.

B. Results

1. Reliability of the Tests

The reliability of the tests (individually) was determined by computing the coefficients of correlation between the combined score in the first and third trials as against that in the second and fourth trials. The average coefficient thus obtained was .91. The actual values are:

Test 1	.94(.01)
" 2	.90(.02)
" 3	.94(.01)
" 5	.96(.01)
" 6	.92(.01)
" 7	.79(.03)

2. Existence of Group Factor of "Speed"

Scores on tests 1, 2, and 7, for the four trials were combined and will be referred to as $speed_1$; similarly the totals on tests 3, 5, and 6 are designated $speed_2$. The sum of the scores in the even-numbered tests in Pintner (g_1), and of the odd-numbered tests in Pintner (g_2), were used in obtaining the following tetrad of correlations:

	$Speed_1$	g_2
$Speed_1$	.51	.05
g_1	.31	.62

The significant size of the tetrad difference (.30, p.e. .07) indicates the presence of a group factor. The division of the speed tests into two pools has given an appreciable "specific correlation," or group factor, which is additional to g.

A closer examination of the tests by means of their intercorrelations has supported the above conclusion. These correlations are represented in Table II, see following page.

From this table an interesting group factor appears between Test 5 and 7. This, however, is in accordance with the findings of several investigators, since both these tests are in the field of arithmetic. The only other suggestion of overlap among the speed

TABLE II

INTERCORRELATIONS BETWEEN THE SEVERAL TESTS

Tests	1	2	3	5	6	7	Pintner
Test 1	.94 (.01)	.35 (.07)	.30 (.08)	.33 (.08)	.26 (.08)	.13 (.08)	.001 (.08)
Test 2	.35 (.07)	.90 (.02)	.11 (.08)	.43 (.07)	.23 (.08)	.27 (.08)	.06 (.08)
Test 3	.30 (.08)	.11 (.08)	.94 (.01)	.15 (.08)	.25 (.08)	.14 (.08)	.25 (.08)
Test 5	.33 (.08)	.43 (.07)	.15 (.08)	.96 (.01)	.17 (.08)	.71 (.04)	.07 (.08)
Test 6	.26 (.08)	.23 (.08)	.25 (.08)	.17 (.08)	.92 (.01)	.16 (.08)	.45 (.07)
Test 7	.13 (.08)	.27 (.08)	.14 (.08)	.71 (.04)	.16 (.08)	.79 (.03)	.16 (.08)

Note:--The Probable Errors in Parentheses.

tests appears in the case of the correlation between Tests 2 and 5. The tetrad differences which include this correlation are not significant, but are suggestive; they may possibly be explained by the fact that both these tests involve numbers. Apart from these two possible exceptions, the correlations of the table seem to indicate some factor operative in the case of all tests. That this factor is not g, is indicated by the extremely low correlations of the tests with Pintner Scores, with the possible exceptions of tests 3 and 6. (Analogies and Word Opposites). These are less mechanical than the other tests; if their correlation with g is eliminated from the correlation that they show with each other. [r(Test 3 vs. Test 6 elim. g) = .16], the equi-proportionality of the coefficients is more marked. We may therefore conclude that the tests here used are related by a common factor other than g. For purposes of reference this is designated as a speed factor.

3. The Relationship Between Speed and Age, With Children in the Same School Grade.

The following table gives the range of chronological age in the two classrooms tested and the mean total speed score corresponding to it. This table shows that there is a definite tendency for the younger children to be superior in the speed score.

This result is illuminated by the following:

(a) The correlation between speed score and intelligence as determined by the Pintner Point Score equals .15 (p.e. .08).

TABLE III

AGE AND CORRESPONDING MEAN TOTAL SPEED SCORE

Age	Frequency	Mean Total Speed Score	Standard Deviation
9	13	859.85	114.5
10	28	827.21	147.
11	17	773.47	115.
12	5	782.6	81.
13	1	692.	-

This again suggests that the "speed" here measured, and intelligence are unrelated.

(b) The I.Q.'s of these children were not available from the results of the Pintner Test, since, although this test is in its final form, mental age equivalents are only now being determined. We were therefore able to use only I.Q.'s that were based on a variety of tests (individual and group), and that had been obtained at different times. Assuming a fair degree of constancy to the I.Q., and that the various intelligence tests gave comparable results, the following data may be quoted as indicative of the relationship between "speed" and I.Q.:

Speed Scores	Mean I.Q.	Mean Pintner Score
High Quartile	104.1 (S.D.6.7)	78.9
Middle 50%	101.6 (S.D.6.5)	70.7
Low Quartile	97.4 (S.D.7.9)	70.5

Since these children are all in the same school grade it appears that the younger ones, who are, on the whole, more intelligent than the older, manifest a greater degree of alertness. A possible cause of this will be discussed later.

4. Relationship Between "Speed" and School Marks.

A composite score representing the school standing of these children was obtained from the teachers. This was based on all important subjects of the school curriculum. Since the teachers knew their own pupils better than those of another class, and since the groups were judged to be equivalent by the Research Division, the following method of arriving at this composite was adopted: Each teacher reduced the scores representing total school standing to a percentage basis. The two distributions thus obtained were equated for means and sigmas.

The school marks correlated with intelligence (Pintner Score) to the degree represented by the coefficient .45 (p.e. .07). The correlation between speed scores and school marks was only .17 (p.e. .08). Averages showed a slight tendency for the speed scores to parallel the school marks; the high quartile middle 50 per cent, and low quartile (speed scores) corresponding to 71.4, 70.7, 69.4, (school marks) respectively.

5. The Comparison Between "Speed Scores" and "Misdemeanors"

The records taken by the Research Division[1] include, for each pupil, the number of misdemeanors throughout the year. A misdemeanor is defined as any act on the part of the child which interfered with classroom procedure. For the year 1930-31, the values that apply to the several speed divisions of the experimental group are as follows:

	No. of Cases	Total Misdemeanors
High Quartile	16	63
Middle 50%	32	158
Low Quartile	16	82

It will be seen that the greater frequency of misdemeanors is somewhat characteristic of the slower pupils, who tend to be, as we have seen, (sect. 3) the older pupils with the lower I.Q.'s.

6. Sex Differences

There do not seem to be any significant sex differences in the speed values obtained. The averages and sigmas for the boys and girls separately are given below:

	No. of Cases	Mean Total Speed Score	Standard Deviation
Boys	34	834.76	139
Girls	30	830.4	147

7. Perseveration and Speed

For one of the classes, scores obtained on a battery of "Perseveration Tests" were available. The correlation between Speed and Perseveration was negligible (-.03). But the following data indicate a trend that might prove to be of interest if examined in more detail:

TABLE IV

TOTAL SPEED SCORES AND MISDEMEANORS FOR PERSEVERATION GROUPS

Basis of Perseveration Scores	No. of Cases	Speed Scores	Misdemeanors (Totals)
First Quartile (Low Perseveration)	7	85.0	3
Middle 50%	14	65.2	70
Fourth Quartile (High Perseveration)	7	70.4	21

This table suggests that while the pupils who manifest least perseveration tend to be among those showing highest speed, the highly perseverative children are not necessarily in the slowest group. In other words, the alertness here, which manifests itself as speed, may overcome to some degree any perseverative tendency present for whatever cause.

8. Increase in Speed Scores in Successive Trials.

The increase in speed scores in successive trials is shown below. It will be seen that there is a fairly consistent gain from trial to trial.

TABLE V

MEAN DOTTING SCORES AND ERRORS FOR EACH TRIAL

1 Trial	2 Mean of Battery of Tests	3 Increase	4 Mean Dotting	5 Mean Errors
1	177.5 (S.D. 28.4)		184.97	12.94
2	193.7 (S.D. 33.2)	16.2	196.08	12.42
3	210.9 (S.D. 39.8)	17.2	197.55	14.16
4	225.2 (S.D. 44.4)	14.3	205.38	14.47

1. The Research Division here referred to, is engaged in studies concerning the behavior of Public School children. This work is conducted by the Toronto Division of the Canadian National Committee for Mental Hygiene, with the cooperation of the Toronto Board of Education. For a description of this work, see Batt, E. A., and Blatz, W. E.: Pedological Seminar and Journal of Genetic Psychology, (December, 1927).

That the scores obtained will depend to some extent upon the degree of motivation in each case, is obvious; but evidence regarding the consistency of motivation is offered in the same Table. Column 4 gives the average score on the Dotting test, which was given as the middle test for each group. It was felt that any marked changes in a subject's attitude, especially in the direction of diminishing interest and keenness, would be indicated by his rate of making a series of dots. Similarly, a change in attitude in the direction of over-keenness would probably result in an increase in the number of errors in the various speed tests. The averages for these errors are given in Column 5. The dotting averages are interpreted as warranting the conclusion that the subjects did not appreciably vary in their attempts to maintain a maximum speed. The errors similarly did not indicate any marked changes, especially when one recalls the fact that the total score on the fourth trial was greater than on earlier trials. An examination of the error records for the subjects individually showed a marked consistency here in that those who made few errors in the first trial did likewise in succeeding trials, while those who made most errors at first did so throughout. In no case was the number of errors significantly large.

The increases from first to fourth trials (total scores) were correlated with intelligence and with speed score Trial One. The values obtained were .10 (p.e. .08) and .41 (p.e. .07), respectively.

In Table VI, Part A, the increases made by the subjects in the first (high) quartile of the distribution of Pintner (intelligence) Scores are compared with those made by the fourth quartile. Similarly for the first and fourth quartiles of the first trial (Part B), and of the total speed scores (Part C).

TABLE VI

INCREASE IN SCORE FOR GROUPS ON BASIS OF PINTNER, FIRST TRIAL AND TOTAL SPEED SCORES

	Trial 1	2	3	4	Incr. 1-2	Incr. 2-3	Incr. 3-4	Incr. 1-4
A. High Quartile (Pintner)	185.00 (26.8)	204.13 (30.5)	221.44 (34.2)	240.5 (35.2)	18.5	17.31	19.69	55.56
Low Quartile (Pintner)	172.31 (31.7)	191.94 (39.0)	206.5 (38.2)	218.25 (44.2)	19.63	15.29	11.13	45.94
Diff. btw. High & Low Quartile	12.69	12.19	14.94	22.25				
B. High Quartile (Trial One)	211.3 (19.9)	237.2 (21.7)	253.2 (30.4)	269.11 (35.2)	25.9	16.	15.91	57.81
Low Quartile (Trial One)	146.4 (11.75)	165.4 (21.0)	179. (26.8)	191.6 (27.4)	19.	13.6	12.6	45.2
Diff. btw. High & Low Quartile	64.9	71.8	74.2	77.51				
C. High Quartile (Total Speed)	210.5 (24.7)	240.88 (24.4)	262.81 (24.1)	280.28 (27.1)	30.28	21.94	17.56	69.88
Low Quartile (Total Speed)	158.56 (14.35)	171.19 (17.9)	177.44 (20.2)	188.38 (16.6)	12.	6.18	10.94	29.19
Diff. btw. High & Low Quartile	51.94	69.69	85.37	91.90				

Note: The (Standard Deviations in parentheses).

The fact that the lowest intelligence quartile gain more from Trial One to Trial Two than do the highest quartile, (Part A), suggests that those subjects of greater intelligence were orientated to the tests more quickly than the others. This advantage was overcome in succeeding trials. This suggestion is further substantiated by the fact that Trials One, Two, and Three correlate with Pintner Scores to the extent .21 (p.e. .08), .19 (p.e. .08), .12 (p.e. .08), respectively. Parts B and C both show that those scoring higher on the speed tests consistently make more improvement than those who score less. These two results are in harmony with the correlations between gain and speed, and gain and intelligence, given earlier.

CONCLUSIONS

1. There is evidence that these tests measure a factor (here called "speed" or "alertness") with a high degree of reliability.
2. This speed appears to be a factor other than g.
3. The "speed" factor is not related to mental age.
4. There is evidence to show that the greater frequency of misdemeanors is somewhat characteristic of the slower pupils.
5. Speed or alertness is not related to "perseveration" but the development of alertness appears to overcome high perseveration.
6. The evidence suggests that "speed" is subject to improvement with practice, and that this improvement does not run parallel to mental age so much as to speed itself.
7. There is some evidence that the speed here measured is in some degree relative to the I.Q. in a classroom situation, since the younger children (those with higher I.Q.) tend to be the most "alert." This may be the result of their competing with older children. The questionable nature of the I.Q. ratings available, however, makes it impossible to be more definite on this point. Further, it is suggested that intensive case studies of individuals are necessary before interpretation of this factor is possible.
8. It is felt that this speed factor may prove to be identified with that mentioned by Kelley (6) and by Hargreaves (4). In the latter case, since Hargreaves interprets his factor as being significant in creative activities, its further exploration might well lead to a method of determining the part played by it in creative work such as composition, and a teacher's evaluation of composition. In this way a possible contribution to education is indicated.

SUMMARY

This study examined intelligence test scores under different time conditions. The degree of gain in scores was relative to the difficulty of the material as well as to the intelligence of the subjects, and some indication was obtained of a group speed factor other than speed of thinking, where easy material was being employed. An attempt was then made to construct tests which depended on this group factor more particularly, and speed tests of a fairly high degree of reliability were constructed. The factor itself was then examined in relation to age, school marks, intelligence, sex, misdemeanors, and perseveration; and possible inferences for education are suggested.

BIBLIOGRAPHY

1. Bernstein, E. Quickness and Intelligence. British Journal of Psychology. Monograph Supplement 1924. (pp. 1-75).
2. Boring and Peak. The Factor of Speed in Intelligence. Journal of Experimental Psychology, 1926. (p. 92).
3. Gilbert, E. C. The Constancy of the Speed Factor in the Performance of Various Tasks. Thesis submitted for Distinction, Harvard University, 1926. Deposited Widener Library, Harvard University. Cited by McFarland, R. A. Role of Speed in Mental Ability. Psychological Bulletin, 1928. (p. 607).
4. Hargreaves, H. L. The "Faculty" of Imagination. British Journal of Psychology. Monograph Supplement, 1927. (pp. 1-74).

5. Highsmith, S. A. The Relation of Rate of Response to Intelligence. Psychological Monograph, 1925. #34. (p. 32).
6. Kelley, T. L. Crossroads in the Mind of Man.
7. May and Terman. The Effect of Doubling the Time Limits in Alpha and Beta Examinations. National Academy of Sciences (p. 416).
8. Reymert. Scandinavian Scientific Review, 1923, 11. Cited by Spearman (5) p. 236.
9. Ruch and Koerth. "Power" vs. "Speed" in Army Alpha Journal Educational Psychology, 1923, (pp. 193-208).
10. Spearman, C. Abilities of Man.
11. Walters, F. C. A Statistical Study of Certain Aspects of the Time Factor in Intelligence. Columbia University Contributions to Education #248. (p. 78).

AN EXPERIMENTAL STUDY OF FUNCTIONAL LEARNING

By A. S. Barr and J. S. Park
University of Wisconsin

The purpose of this experiment was to study the relative amounts of learning procured by subjects employing the so-called direct and incidental methods of instruction. No attempt will be made to define direct and incidental instruction since these terms are employed with various meanings under different times, places, and conditions. Instead, the writers start with two carefully defined methods of study which may or may not constitute direct and incidental instruction according to the reader's definition. The results are true only for the subjects, methods, materials, conditions, and the learning measured in this experiment.

An attempt will be made to answer the following questions:

1. When two equally difficult artificial alphabets are studied under the conditions of this experiment, and the results measured by direct recall and recognition tests such as those employed in this investigation, by which of the two methods of study employed in this investigation, does the pupil make the greater gross gain in raw score?
2. What are the differences in the shapes of the learning curves for the two methods of study employed in this investigation?
3. What is the practice effect of learning by one method, on learning by the other?
4. How do the learning curves for the two methods of study compare when practice effect is held constant?
5. What is the relationship between the number of symbols practiced, during each study period, and the number of symbols learned, for each method?
6. How do the results from the recall test and the recognition test agree for the two methods of learning?
7. What is the relationship between the number of symbols learned by the indirect method and the knowledge of the content of the material translated?
8. What is the relationship between the number of symbols learned and the intelligence of the subjects?
9. What differences in attitude, effort, and fatigue did the subjects observe in the use of the two methods of study?

Two artificial alphabets were employed as the subject matter of this experiment. In order to make the two alphabets of equal difficulty, they were constructed through the use of squares, circles, angles, triangles, etc., upon identical patterns. The two alphabets are given below:

ALPHABET A

a, △3; b, □1; c, ④; d, ◇2; e, □5; f, 3; g, △1; h, □2; i, 1; j, /5\; k, ◇1; l, ⑤; m, 2; n, ▽4; o, □2; p, /1\; q, □3; r, ①; s, ◇5; t, 4; u, □1; v, ▽5; w, □4; x, ▽2; y, 5; z, \1/.

ALPHABET B

a, ②; b, ▽3; c, □1; d, □4; e, ◇1; f, □3; g, 1; h, □2; i, □4; j, ▽1; k, 5; l, /4\; m, \2/; n, 3; o, /2\; p, □5; q, △2; r, ◇4; s, □2; t, /3\; u, 4; v, \3/; w, △4; x, 2; y, ③; z, \5/.

Two methods of study were used. By one method (A), the direct method, the subjects were instructed to concentrate upon the memorization of the symbols, while by the other method (B), the incidental method, the subjects were instructed to concentrate upon the translation of meaningful material, the content of which they were to be tested later. The following directions were read to

the subjects after the practice materials had been placed in their hands:

DIRECTIONS READ TO BOTH GROUPS A AND B

This is an experiment in learning, and is not an intelligence test. The object of the experiment is to compare the learning curves obtained during the acquisition of an artificial alphabet, by two different methods. One method, A, is by direct memorization of the symbols by purposeful learning, while the other method, B, is by incidentally learning the symbols while translating meaningful material. The time intervals for each learning period will be 9 minutes. The testing time will be 10 minutes, divided into three tests of 2, 5, and 3 minutes respectively. There will be four practice periods, each followed by a testing period, e.g. Practice, test, practice, test, etc., for four practice periods.

DIRECTIONS READ TO GROUP A (DIRECT)

When the signal is given you will have 9 minutes to study an artificial alphabet. Start with the letter "a" and write the symbol for each letter three times on the dotted line following that letter. Go through the entire alphabet. Then turn the page and try to recall the symbols from memory. Write the symbol for each letter three times. When you cannot recall the symbol quickly, turn back to page 1 and refer tc the key. Then write the symbol three times and put a check mark (✓) after the symbol. Do the same for each page of the practice booklet. (Each subject was provided with 4 practice booklets, one to be used during each study period.)

DIRECTIONS READ TO GROUP B (INCIDENTAL)

Remove the top page of the pad and place it on the bottom of your pad with its left margin protruding far enough to permit the key to be seen readily for convenient reference in translating the practice material. Start at the beginning of page 1 and transcribe every letter into its appropriate symbol. Make no conscious effort to memorize the symbols beyond the incidental learning which occurs spontaneously and automatically. The purpose of the experiment is to compare the average performance of each group as a study of method, and is not a study of individual performance. Work as fast as you can, for the number of letters translated is to be counted as a part of the score to be compared with the number of letters written by the group using the direct method. Start at the beginning of each page and go as far as you can in the nine minutes allowed. (The subjects were provided with appropriate practice booklets.)

Directions similar to those read audibly were printed on the cover page of each booklet. The following is a sample page of the practice material for method B.

SAMPLE OF PRACTICE MATERIAL
FROM BOOKLET B

1. B e n i g n = k i n d ‖ c h u r l i s h
2. W h e e d l e = c a j o l e ‖ r e b u k e
3. S u a v e = u n c t u o u s ‖ b r u s q u e
4. F r a t e r n i z e = s y m p a t h i z e ‖ o s t r a c i z e
5. P r o l i x = d i f f u s e ‖ c u r t
6. E x c r u c i a t e = a g o n i z e ‖ t r a n q u i l i z e
7. W e a k n e s s = d e b i l i t y ‖ v i g o r
8. A z u r e = c e r u l e a n
9. W r y = s k e w ‖ s y m m e t r i c
10. O b n o x i o u s = p e r n i c i o u s

Note: Words connected by "=" mean the same thing; words connected by "‖" are opposites.

The rotated group plan of experiment was adopted. The subjects of this investigation were graduate students enrolled in an advance course in Experimental Education (Education 208) at the University of Wisconsin. Two groups of twenty-six subjects each were used, one group being given the direct method of study (A) first, while the other group was given the incidental method of instruction (B) first. Three weeks later, allowing for sufficient time to prevent interference, the methods were reversed, and method B was given to those who had had method A, and method A was given to those subjects who had had method B first. In this way each subject served as his own control. A retention test was

given to each group one week after each experimental learning period. In making the comparisons, the results for the first applications of methods A and B, the second applications of methods A and B, the first and second applications of method A, the first and second applications of method B, and the combined applications of method A with the combined applications of method B, were all compared. In this way the effect of the practice of having taken the alternate method previously could be measured and held constant.

Three tests were employed to measure the amount of learning secured from the two methods of study and the knowledge of the content of the material being translated: test 1, a direct recall test of the symbols learned; test 2, a matching test of the synonyms and antonyms translated into the B alphabet; and test 3, a recognition test of the symbols learned. These tests are reproduced below:

TEST 1 - DIRECT RECALL

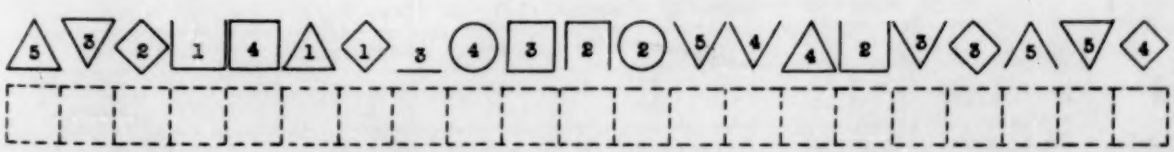
etc., for the entire alphabet.

TEST 2 - MATCHING TEST

MATCH THE ITEMS IN THESE COLUMNS WITH THOSE IN THE TWO OPPOSITE COLUMNS

		SAME	OPPOSITE
1 Abstemious	21 Juxtaposition	( ) Agile	( ) Believer
2 Affluence	22 Knack	( ) Agonize	( ) Blandiloquence
3 Agnostic	23 Levity	( ) Azure	( ) Brusque
4 Ascititious	24 Maze	( ) Bonny	( ) Churlish
5 Bamboozle	25 Obnoxious	( ) Cajole	( ) Dirge
6 Benign	26 Paean	( ) Diffuse	( ) Disrelish
7 Breeze	27 Piquant	( ) Elicit	( ) Exsiccation
8 Buxom	28 Prize	( ) Excrescence	( ) Forfeiture
9 Castigate	29 Prolix	( ) Flagellate	( ) Frank
10 Cerulean	30 Skew	( ) Frivolity	( ) Germane
11 Debility	31 Spry	( ) Frugal	( ) Graceful
12 Educe	32 Subtle	( ) Hoax	( ) Ineptitude
13 Equivocal	33 Superfluity	( ) Labyrinth	( ) Insipid
14 Excruciate	34 Unctuous	( ) Luxury	( ) Lugubrious
15 Exhalation	35 Vanquish	( ) Nefarious	( ) Melancholy
16 Facetious	36 Vociferate	( ) Perplexing	( ) Ostracize
17 Fraternize	37 Wicked	( ) Pernicious	( ) Panegyrize
18 Gawky	38 Wheedle	( ) Propinquity	( ) Succumb
19 Harangue	39 Zealot	( ) Pungent	( ) Symmetric
20 Jocund	40 Zest	( ) Visionary	( ) Tranquilize
		( ) Zephyr	( ) Vigor

TEST 3 - RECOGNITION



and so on for the entire list of symbols used.

TABLE I

THE GROSS NUMBER OF LETTERS MEMORIZED DURING FOUR STUDY PERIODS OF NINE MINUTES EACH, AND THE NUMBER RETAINED AFTER ONE WEEK

Measures	No. symbols learned in 36 minutes				No. symbols retained after one week			
	No. Correct		No. Errors		No. Correct		No. Errors	
	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m
Method A	18.37	.78	1.48	.22	4.68	1.20	3.05	.47
Method B	11.69	.78	1.38	.17	1.58	.51	3.00	.65
Difference	6.68	1.10	.10	.28	3.10	1.58	.05	.80
Critical Ratio	6.06		.36		1.96		.06	

A comparison will first be made of the gross number of letters learned by the two methods of study. The pupils employing the direct method learned 18.37 out of the 26 in the alphabet in 36 minutes, and after one week could recall 4.68 letters, while the pupils employing the incidental method learned 11.69 letters and retained only 1.58 letters. The differences are statistically significant. The pupils employing the direct method of learning retained approximately 25 per cent of the materials learned, while those employing the incidental method retained approximately 13 per cent. The differences between the error scores are not significant statistically.

TABLE II

A COMPARISON OF THE LEARNING CURVES FOR DIRECT AND INCIDENTAL LEARNING

No. of correct symbols, as measured by a direct recall test (Test 1)

Measures	Period 1		Period 2		Period 3		Period 4		Retention		Total	
	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m
Method A	4.29	.46	9.52	.63	14.27	.81	18.37	.78	4.68	1.20	46.25	2.41
Method B	3.83	.35	6.83	.49	9.38	.71	11.69	.78	1.58	.51	3..73	2.18
Difference	.46	.90	2.69	.80	4.89	1.08	6.68	1.10	3.10	1.58	14.52	3.25
Critical Ratio	.51		3.37		4.54		6.06		1.96		4.47	

No. of errors as measured by direct recall (Test 1)

	1		2		3		4		Retention		Total	
Method A	2.71	.27	2.10	.26	1.71	.24	1.48	.22	3.05	.47	8.00	.66
Method B	.96	.13	1.19	.16	1.44	.20	1.38	.17	3.00	.65	4.88	.45
Difference	1.75	.30	.91	.31	.27	.31	.10	.28	.05	.80	3.12	.80
Critical Ratio	5.84		2.98		.87		.36		.06		3.90	

No. of letters correct as measured by a recognition test (Test 3)

	1		2		3		4		Retention		Total	
Method A	3.06	.33	7.21	.56	10.65	.74	13.37	.73	4.95	.77	34.08	2.09
Method B	3.00	.30	5.60	.38	7.56	.58	10.54	.73	1.77	.41	26.69	1.82
Difference	.06	.45	1.61	.67	3.09	.94	2.83	1.03	3.18	.87	7.39	2.77
Critical Ratio	.13		2.40		3.29		2.75		3.66		2.67	

No. of errors as measured by a recognition test (Test 3)

	1	2	3	4	Retention	Total	
Method A	2.69	2.21	1.73	1.81	4.90	8.40	.72
Method B	1.19	1.40	1.52	1.54	4.0	5.67	.58
Difference	1.50	.81	.21	.27	.90	2.73	.93
Critical Ratio						2.94	

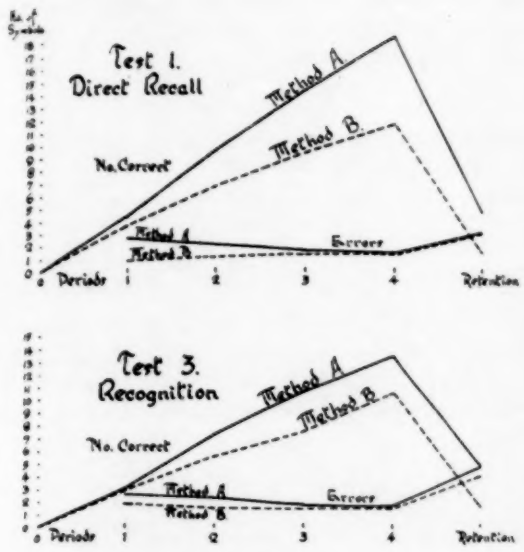


Figure 1.

The learning curve for each method is approximately a straight line, the increments of learning for each period being approximately equal. The learning process was not carried far enough by either method to show any leveling off of the curves. The curve for method A shows the faintest suspicion of a positively accelerated curve at the start, the acceleration becoming negative by the fourth period. The curve for method B shows a faint negative acceleration from the start, becoming more evident by the fourth period. The differences between the curves becomes greater with each learning period, the critical ratios from the second period on being statistically significant.

There is a considerable difference in favor of the second application, in the learning curves for both methods of study. In general, there is, however, a gradual decrease in the difference from the first to the fourth study period, indicating that the effect of practice is gradually lost as study advances. This difference is more marked in the case of method A. It would seem from these data that the experience of the previous application increases the scores for both methods of study, but that this practice effect gradually disappears as learning progresses. One interesting sidelight is that while the advantage is in favor of the second application throughout both series, in the retention tests, the scores favor the first application.

In order to rule out or hold constant the effect of practice in the learning of the artificial alphabet, the mean scores for each method were compared within the same application. This amounts to treating the experiment as consisting of two pairs of equivalent groups. While the means for the second application due to the effect of practice are higher throughout the series for both methods of study than for the first application, the differences between the results for the two methods are practically the same, the critical

TABLE III

A COMPARISON OF THE LEARNING CURVES TO DETERMINE THE PRACTICE EFFECT OF EACH METHOD OF LEARNING ON THE OTHER

Measures	Period 1		Period 2		Period 3		Period 4		Retention		Total	
	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m
Method A												
Application 1	3.35	.51	8.42	.72	13.73	1.08	18.27	1.07	6.62	1.60	43.73	2.99
Application 2	5.23	.72	10.62	.98	14.81	1.19	18.46	1.16	1.89	1.33	48.77	3.70
Difference	1.88	.88	2.20	1.22	1.08	1.61	.19	1.58	4.73	2.08	5.04	4.76
Critical Ratio	2.13		1.81		.67		.12		2.27		1.06	
Method B												
Application 1	3.00	.49	5.92	.76	8.54	1.08	10.50	1.21	2.53	.68	27.96	3.37
Application 2	4.65	.45	7.73	.57	10.23	.90	12.89	.92	.27	.13	35.50	2.58
Difference	1.65	.67	1.81	.95	1.69	1.41	2.39	1.52	2.26	.69	7.54	4.24
Critical Ratio	2.46		1.91		1.20		1.57		3.28		1.78	

ratios being almost identical for the two applications. When compared with the means of the two applications combined, the differences are practically the same, but the critical ratios for the latter are higher than for either application alone, due to the higher statistical reliability of the larger rotated group.

TABLE IV

A COMPARISON OF THE LEARNING CURVES FOR THE TWO METHODS OF STUDY WHEN THE EFFECT OF PRACTICE IN THE PREVIOUS APPLICATION IS HELD CONSTANT

	Period 1		Period 2		Period 3		Period 4		Retention		Total	
	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m
Application 1												
Method A	3.35	.51	8.42	.72	13.73	1.08	18.27	1.07	6.62	1.60	43.73	2.90
Method B	3.00	.49	5.92	.76	8.54	1.08	10.50	1.21	2.53	.68	27.96	3.39
Difference	.35	.71	2.50	1.05	5.19	1.53	7.77	1.62	4.09	1.74	15.77	4.51
Critical Ratio	.49		2.39		3.39		4.81		2.35		3.39	
Application 2												
Method A	5.23	.72	10.62	.98	14.81	1.19	18.46	1.16	1.89	1.33	48.77	3.70
Method B	4.65	.45	7.73	.57	10.23	.90	12.89	.92	.27	.13	35.50	2.58
Difference	.58	.85	2.89	1.13	4.58	1.49	5.57	1.48	1.62	1.34	13.27	4.51
Critical Ratio	.68		2.55		3.07		3.76		1.21		2.94	

TABLE V

A COMPARISON OF THE AMOUNT LEARNED DURING EACH PERIOD WITH THE AMOUNT OF PRACTICE MATERIAL STUDIED DURING THAT PERIOD

	Period 1		Period 2		Period 3		Period 4		Total
Measures	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m	
Method A, Application 1									
No. Symbols Written	46.23	2.00	48.65		57.65		63.46		
Accumulated Sum	46.23	2.00	94.88	4.26	152.54	6.88	216.00	9.27	
Correlation with No. Symbols Learned	.075		.037		.026		.109		.174
Method B, Application 2									
No Symbols written	53.62	2.92	60.92		70.31		76.44		
Accumulated Sum	53.62	2.92	114.56	5.68	184.85	9.10	258.35	10.01	
Correlation with No. Symbols Learned	.548		.488		.613		.625		.629
Method A, two Applications Combined No. Symbols Written	49.92	1.84	54.79		63.96		69.82		
Accumulated Sum	49.92	1.84	104.71	3.80	168.69	6.13	237.17	7.43	
Correlation with No. Symbols Learned	.442		.385		.378		.362		.447
Method B, Application 1	123.00	6.96	118.88		137.65		150.92		
Accumulated Sum	123.00	6.96	237.92	14.32	375.58	21.80	526.50	29.57	
Correlation with No. Symbols Learned	.307		.277		.224		.324		.270
Method B, Application a									
No. Symbols Written	126.23	6.10	128.85		131.65		138.36		
Accumulated Sum	126.23	6.10	255.08	11.22	386.73	16.04	521.35	21.33	
Correlation with No. Symbols Learned	.337		.237		.172		.024		.146
Method B, two Applications Combined No. Symbols Written	124.65	4.62	123.87		134.65		141.98		
Accumulated Sum	124.65	4.62	246.50	9.17	381.15	13.55	523.92	18.23	
Correlated with No. Symbols Learned	.332		.284		.210		.208		.214

TABLE V (Continued)

Correlation of Practice Material A with Practice Material B

	Method A, Application 1	Application 2	Two Applications Combined
Method B, Application 1		.753	
Method B, Application 2	.444		
Method B, Two Applications combined			.435

In order to determine the correlation between the number of symbols learned and the number of symbols practiced during each study period, a count was made of the number of symbols written during each study period. These are presented in Table V above. The relatively low coefficients of correlation throughout would seem to indicate that frequency of repetition alone does not adequately account for the learning secured. In method A there is a gradual increase in the number of symbols written from one period to the next throughout the series, while in method B there is a decrease in the number written during the second period from the number written during the first period. This would seem to indicate that for method B, after the first testing period, the subjects began to mentally rehearse the meaning of the symbols at the expense of writing them. The correlation between the amount of practice material written for methods A and B would indicate that there was more in common in the writing of the symbols for the two methods of study, than between the memorization of the symbols and the writing of them by the same method of study.

TABLE VI

A COMPARISON OF TWO WAYS OF TESTING (TEST 1, RECALL; TEST 3, RECOGNITION) AND THE SAME AMOUNT OF LEARNING

	Period 1		Period 2		Period 3		Period 4		Retention		Total	
Measures	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m
Method A, Diff. between Tests 1 and 3, Application 1	.58	.64	2.19	.91	3.46	1.42	5.31	1.34	.21	2.18	11.50	3.79
Critical Ratio	.91		2.41		2.44		3.96		.10		3.03	
Correlation	.790		.853		.805		.837		---		.887	
Diff. between Tests 1 and 3, Application 2	1.88	.89	2.43	1.34	3.77	1.64	4.69	1.68	.55	1.81	12.85	5.04
Critical Ratio	2.11		1.81		2.30		2.79		.30		2.55	
Correlation	.831		.953		.840		.828		---		.948	
Diff. between Tests 1 and 3, Combined Applications	1.23	.57	2.31	.84	3.62	1.10	5.00	1.14	.27	1.43	12.17	3.19
Critical Ratio	2.16		2.75		3.29		4.39		.19		3.82	
Correlation	.812		.924		.823		.820		---		.924	
Method B, Diff. between Test 1 and 3, Application 1	.85	.61	1.23	.94	1.69	1.29	1.08	1.67	.20	.89	4.84	4.36
Critical Ratio	1.39		1.31		1.31		.65		.22		1.11	
Correlation	.838		.831		.884		.874		---		.962	
Diff. between Tests 1 and 3, Application 2	.80	.62	1.23	.73	1.92	1.14	1.24	1.26	.19	.24	5.23	3.35
Critical Ratio	1.29		1.68		1.68		.98		.79		1.56	
Correlation	.846		.792		.870		.924		---		.934	
Diff. between Tests 1 and 3, Combined Appl.	.83	.46	1.23	.62	1.82	.92	1.15	1.07	.19	.65	5.04	2.84
Critical Ratio	1.80		1.98		1.98		1.07		.29		1.77	
Correlation	.831		.828		.881		.897		---		.954	

From these data it is apparent that the recall method of testing the learning of an artificial alphabet results in higher scores than the recognition method of testing. This result is somewhat different than one would ordinarily expect. It was also thought that one form of testing might favor one or the other method of study. While the results for the recall test and method A are higher than those for method B, the differences are approximately proportional to the means from which they are taken. It would seem that both tests measure to a large extent, the same thing, inasmuch as the correlations (practically all in the .80's and .90's) between these two methods of testing are so high. The recall method gives a smoother and more uniform curve than the recognition method of testing.

TABLE VIII

A COMPARISON OF THE INTELLIGENCE OF THE SUBJECTS AND THE AMOUNT LEARNED

Groups	Intelligence		Total Test 1		Correlation		
	Mean	σ m	Mean	σ m	r	σ	N
Method A, Application 1	228.00	17.69	39.25	3.91	.174	.280	12
Method A, Application 2	233.15	14.11	55.08	5.11	.214	.265	13
Method A, Applications 1 and 2 combined	230.68	11.23	47.48	3.62	.191	.193	25
Method B, Application 1	233.15	14.11	31.23	4.98	.351	.243	13
Method B, Application 2	228.00	17.69	38.67	3.99	.114	.285	12
Method B, Applications 1 & 2 combined	230.68	11.23	34.80	3.31	.120	.197	25

TABLE VII

A COMPARISON OF THE NUMBER OF SYMBOLS LEARNED WITH THE KNOWLEDGE OF THE CONTENT OF THE MATERIAL TRANSLATED DURING THE INCIDENTAL LEARNING

	Number of Words Correct								
Measures	Period 1		Period 2		Period 3		Period 4		Correlation
	Mean	σ m	Mean	σ m	Mean	σ m	Mean	σ m	r
Method B, Application 1	5.11	6.27	8.08	9.12	11.80	8.47	34.81	5.28	.503
Method B, " 2	4.76	5.19	7.88	8.54	10.31	8.42	31.92	3.51	.032
Difference	.35	1.08	.20	.58	1.49	.05	2.89	6.34	
Critical Ratio	*	*	*	*	*	*	.46		
Method B, Applications 1 and 2 Combined	4.91	5.73	7.98	8.83	11.04	8.44	33.37	2.53	.391
Increment		.82	2.25	.85	2.21				

A matching test was devised for measuring the knowledge of the content of the material translated by method B. The directions for method B emphasized concentration upon the translation of the text material, rather than the learning of the symbols translated. The content of the material to be translated was purposely made difficult, in an effort to find words not already familiar to the subjects, so that the learning measured would be true learning. The correlation between the number of symbols used and the knowledge of the content of the material translated being .391 indicates a fairly low relationship. This may mean that the material was too difficult in nature, or that the subjects concentrated more on the actual translation than on the learning of the meaning of the words being used.

The intelligence ratings on the American Council Psychological Examination were available for about half of the entire group only. The mean score for one group was 228, which fell in the 92nd percentile and for the other group was 233, which fell in 93rd percentile, so that the two groups were approximately equal in intelligence as measured by the American Council Psychological Examination. The correlation between intelligence and the learning of the symbols was .191 for the direct method of learning, and .120 for the incidental method, which seems to indicate that there is very little if any relationship between intelligence and the ability to master an artificial alphabet under the conditions of this experiment.

TABLE IX

A COMPARISON OF THE EFFECTS ATTITUDE, EFFORT, AND FATIGUE ON THE LEARNING OF AN ARTIFICIAL ALPHABET, BY TWO DIFFERENT METHODS

	Attitude		Effort		Fatigue	
	Mean	σ m	Mean	σ m	Mean	σ m
Method A, Application 1	2.04	.29	2.31	.23	3.15	.17
Method A, Application 2	2.00	.23	2.04	.19	3.42	.20
Difference	.04	.37	.27	.30	.27	.26
Critical Ratio	.11		.90		1.04	
Method B, Application 1	2.81	.25	2.23	.23	3.50	.24
Method B, Application 2	1.92	.20	2.42	.25	3.38	.24
Difference	.89	.32	.19	.34	.12	.34
Critical Ratio	2.72		.56		.35	
Method A, Combined 1 & 2	2.02	.18	2.17	.15	3.29	.13
Method B, Combined 1 & 2	2.37	.17	2.33	.17	3.44	.17
Difference	.35	.25	.16	.23	.15	.21
Critical Ratio	1.40		.70		.71	

A further study was made of the relation between attitude, effort, and fatigue, and the learning by the two methods of study. These data are presented in Table IX. The differences all appear to be small and not significant statistically. Under the conditions of this experiment these data would seem to indicate: (1) that, the effort put into, and the attitude toward the experiment were considerably better than average, (3 being average on a 5 point scale, 1 being very favorable and 5 a very unfavorable attitude, 1 being great effort and 5 little or no effort, 1 no fatigue and 5 much fatigue), but that fatigue was greater than average; (2) that attitude and effort ranked slightly better for method A than for B, but that fatigue was somewhat greater for method B than for A; (3) that the attitude was better toward the experiment during the second application than during the first, for both methods; (4) that there was more effort put forth in the second application of method A and the first application of method B; and (5) that fatigue was greater for method A in the second application, and for method B in the first application.

SUMMARY AND CONCLUSION

In conclusion, the data of this experiment would seem to indicate that when two equally difficult artificial alphabets are learned by direct memorization and by incidental learning, the following are true:

1. Direct learning is about 57 per cent more efficient than incidental learning in the study of two equally difficult artificial alphabets, the difference in favor of direct learning being 6 times its standard error.

2. There is a very marked difference in the general trend of the learning curve for each method of study. The differences in the two curves, measured after each study period, are three or more times their standard error for every period beyond the first.

3. There is a considerable practice effect from one method to the other, resulting from having studied the equivalent alphabet previously.

4. When practice effect is held constant, the results of this experiment show that direct memorization is more effective in learning an artificial alphabet than incidental learning, by 74 per cent for the first application and 43 per cent on the second application, the critical ratios being approximately 4 in both cases.

5. There is a moderately low correlation (.447 for method A and .214 for B) between the number of symbols learned and the number of symbols written in practice.

6. The recall method of testing the learning of the alphabets, shows a more consistent curve than the recognition method of testing for this type of material.

7. There is a moderately low correlation (.391) between the number of symbols learned and the knowledge of the content of the material translated by the incidental method of learning.

8. There is a very low correlation (.191 and .120 for methods A and B respectively) between the intelligence of the subjects and the number of symbols learned by either method.

9. The incidental method of learning an artificial alphabet was found to be more fatiguing than direct memorization, under the conditions of this experiment.

The reader is again warned that these results are true only for the subjects, materials, tests and conditions of this experiment. One of the greatest limitations of the experiment was probably the absence of a truly incidental learning situation.

THE RELATIVE CONTRIBUTION OF CERTAIN FACTORS TO INDIVIDUAL DIFFERENCES IN ARITHMETICAL PROBLEM SOLVING ABILITY[1]

By Max D. Engelhart
University of Illinois

The problem of this investigation is to determine the relative contributions of intelligence, computation ability, and reading ability to individual differences in arithmetical problem solving ability. It may be restated more precisely as follows: What per cents of the variance[2] of arithmetical problem solving ability of fifth-grade pupils as measured by Test 1 of the New Stanford Arithmetic Test, Form X, are due to variation in intelligence as measured by the Otis Self-Administering Test of Mental Ability, Intermediate Examination, Form C, to variation in reading ability as measured by the New Stanford Reading Test, Form X, and to variation in computational ability as measured by Test 2 of the New Stanford Arithmetic Test, Form X?

This problem is not an academic one. Achievement in a given field is influenced by several factors. In planning where the emphasis shall be placed in classroom instruction, it is helpful to know to what factors attention should be given. For example, if it is found that achievement in history is significantly affected by achievement in reading it may be concluded with considerable justification that instructing pupils in the reading of history is desirable. Solution of the problem stated above should be helpful in drawing inferences with respect to the types of instruction which are likely to result in increased arithmetical problem solving achievement.

The variables in this study are referred to frequently as intelligence, reading ability, computation ability, and problem solving ability. Furthermore, the terms ability and achievement are used synonymously. More precisely the variables are the scores made on the tests administered, but it would be awkward to use this phraseology consistently.

The data used in this study are the scores of 568 fifth-grade pupils of the Decatur, Illinois, Public Schools. Tables I to IV give the means, standard deviations, and basic correlations. The numbers assigned to the variables in Table I are used as symbols in the other tables.

TABLE I

MEASURES OF CENTRAL TENDENCY AND VARIABILITY

	Mean	Standard Deviation
1 Problem score	67.3770	14.9325
2 Intelligence score	27.8785	10.2860
3 Computation score	62.8345	8.7045
4 Reading score	69.4365	12.8145
5 Chronological age	127.2682 Mo.	11.0829 Mo.

In connection with Table II it may be observed that the probable errors of these coefficients are sufficiently small in magnitude to make all of the coefficients significant. The smallest coefficient is more than ten times its probable error. The correlations with chronological age are low since the group consisted of pupils in a single grade. (See following page for Table II).

In Table III, following page, the zero order coefficients of correlation have been corrected for attenuation by means of the following formula:

$$r_{12} = \frac{r_{12}}{\sqrt{r_{11}\ r_{22}}}$$

1. The writer wishes to express his appreciation to Superintendent William Harris, Supervisor Sarah Mark Imboden, and numerous teachers of the Decatur, Illinois, Public Schools for administering and scoring the tests used in this study. He wishes to express his thanks to Dr. W. S. Monroe of the Bureau of Educational Research, University of Illinois, to Dr. Barbara Stoddard Burks of the Pasadena Public Schools, and to Dr. Bradford B. Smith of the Cleveland Trust Company, Cleveland, Ohio, for valuable suggestions while the research was in progress and for critical reading of the completed report of the research. Finally, mention should be made of the aid rendered in checking scores and in computation by Mrs. George Curzon, a clerk of the Bureau of Educational Research.

2. "Variance" is defined as the square of the standard deviation. For reasons for using this measure of variability, see:

Burks, B. S. "The Relative Influence of Nature and Nurture Upon Mental Development; A Comparative Study of Foster Parent--Foster Child Resemblance and True Parent--True Child Resemblance," The Twenty-Seventh Yearbook of the National Society for the Study of Education, Part I (Bloomington, Illinois: Public School Publishing Company, 1928), p. 298.

TABLE II

ZERO ORDER COEFFICIENTS OF CORRELATION WITH THEIR PROBABLE ERRORS

	2	3	4	5
1	.5898 ± .0184	.5821 ± .0187	.5008 ± .0212	-.3182 ± .0254
2		.4946 ± .0214	.7878 ± .0107	-.3222 ± .0254
3			.4378 ± .0229	-.3013 ± .0257
4				-.2794 ± .0261

Each of the reliabilities reported in the test manuals was made to conform to the variability of the group of pupils used in this study by means of Kelley's formula for the relation between ranges in obtained scores and reliability coefficients.[1]

TABLE III

ZERO ORDER COEFFICIENTS CORRECTED FOR ATTENUATION

	2	3	4	5
1	.7019	.8091	.5817	-.3459
2		.6926	.9218	-.3528
3			.5983	-.3853
4				-.2986

In Table IV coefficients are given for the correlations between the different variables with the variable chronological age held constant. The ordinary partial correlation formula for three variables was used with the coefficients corrected for attenuation.

No correlation technique can show the direction of causation. Which of two or more variables are causes and which are effects must be determined by logical analysis or by experiment. In the present study, it is assumed <u>a priori</u> that problem solving achievement is contributed to by the variables intelligence, computation achievement, reading achievement, and unknown causes. This seems to be a reasonable assumption.

TABLE IV

COEFFICIENTS CORRECTED FOR ATTENUATION AND WITH CHRONOLOGICAL AGE CONSTANT

	2	3	4
1	.6605	.7806	.5343
2		.6447	.9143
3			.5487

The coefficient of correlation .6605 is not an index of the contribution of intelligence to problem solving ability, because the group is heterogeneous with respect to computation ability and reading ability. The coefficient of correlation .7806 is not an index of the contribution of computation since the group is heterogeneous with respect to intelligence and reading ability. The coefficient of correlation .5343 is not an index of the contribution of reading ability since the group is

1. Kelley, T. L. <u>Statistical Method</u> (New York: The Macmillan Company, 1923), p. 222.
The reported reliability of the arithmetic reasoning test is .75 for $\sigma = 11.7$; hence, $r_{11} = .8465$ for $\sigma_1 = 14.9325$.
The reported reliability of the Otis test is .948 for $\sigma = 18.38$ (obtained from the reported probable errors of a score); hence, $r_{2II} = .8340$ for $\sigma_2 = 10.2860$.
The reported reliability of the computation test is .86 for $\sigma = 14.5$; hence $r_{3III} = .6115$ for $\sigma_3 = 8.7045$.
The reported reliability of the reading test is .872. Since standard deviations are reported only for the sub-tests in word meaning and paragraph meaning 12.9 and 14.1, the standard deviation corresponding to .872 was estimated as follows: The correlation between the two sub-tests was assumed to approximate .75. The standard deviation needed was then secured by the formula:

$$\sigma_{sum} = \sqrt{12.9^2 + 14.1^2 + 2 \cdot .75 \cdot 12.9 \cdot 14.1}$$

The resulting σ was divided by two since the scores of the two sub-tests are averaged. It equals 12.63. Hence, $r_{4IV} = .8757$ for $\sigma_4 = 12.8145$.
The reliability of the chronological ages was assumed to be 1.00.

heterogeneous with respect to intelligence and computation ability. Partial correlation is often used to eliminate such heterogeneity. This technique, however, partials out too much when one or both of the variables whose net correlation is to be determined are not merely effects of the variables eliminated, or held constant.[1] For example, in the problem of the relative contributions of parental intelligence and environment to child intelligence the use of partial correlation to eliminate or hold constant environment results in the elimination of too much. Parental intelligence affects environment. When environment is partialled out the net correlation between parental intelligence and child intelligence is too low. When environment is partialled out, all that part of parental intelligence which correlates with environment is removed, and this probably contains a significant proportion of the hereditary contribution. Partial correlation techniques are likewise inadequate for the present problem since each of the variables intelligence, computation achievement, and reading achievement are both causes and effects of each other. The net correlations of each of them with problem solving ability, obtained by the partial correlation techniques, would be too low. For example, when intelligence and reading ability are partialled out, all that part of computation ability which correlates with intelligence and reading ability is removed, and this probably contains a significant proportion of the computation ability contribution.

The path coefficient technique provides for the measurement of both direct and indirect influences.[2] Its use involves the calculation of coefficients of determination.[3] Some of these, when the decimal point is moved two places to the right, show the per cents of variance of the dependent variable which are due to direct influences from the several independent variables, or causes. Others of the coefficients of determination show what are assumed to be the combined, or joint, contributions of given pairs of the independent variables, over and above the direct contribution of each, to the variance of the dependent variable. When these coefficients of determination are separated according to the weights indicated by the coefficients of direct determination, the assumed indirect influences of each of the two variables of the pair are given. The per cents which represent the indirect influences can then be combined with the per cents which indicate the magnitudes of the direct influences. Hence, the analysis finally results in a per cent for each of the independent variables which represents its total contribution to the variance of the dependent variable. When these per cents are added together they stand for the total contribution of all the known independent variables to the variance of the dependent variable.

---

1. Burks, B. S. "On the Inadequacy of the Partial and Multiple Correlation Technique," Journal of Educational Psychology, (November and December, 1926) XVII, 532-40; 625-30.
   Burks, B. S. "Statistical Hazards in Nature-Nurture Investigations," Twenty-Seventh Yearbook of the National Society for the Study of Education, Part I (Bloomington, Illinois: Public School Publishing Company, 1928), pp. 9-33.
2. The path coefficient technique was contributed by Dr. Sewall Wright of the University of Chicago (formerly of the Bureau of Animal Industry, United States Department of Agriculture).
   Wright, Sewall. "Correlation and Causation," Journal of Agricultural Research, XX (January, 1921), 557-85.
   Wright, Sewall. "The Theory of Path Coefficients," Genetics, VIII, (May, 1923) 238-55.
   Two previous studies in the field of education in which the techniques of path coefficients were used were helpful in planning the present study:
   Burks, B. S. "The Relative Influence of Nature and Nurture Upon Mental Development," The Twenty-Seventh Yearbook of the National Society for the Study of Education, Part I (Bloomington, Illinois: Public School Publishing Company, 1928), pp. 299-304.
   Heilman, J. D. "The Relative Influence Upon Educational Achievement of Some Hereditary and Environmental Factors," The Twenty-Seventh Yearbook of the National Society for the Study of Education, Part II (Bloomington, Illinois: Public School Publishing Company, 1928), pp. 35-65. For a more extended account, see:
   Heilman, J. D. "Factors Determining Achievement and Grade Location," Journal of Genetic Psychology, XXXVI (September, 1929), 435-56.
3. A coefficient of determination when the decimal point is shifted two places to the right gives the per cent of variance in a dependent variable which is due to an independent variable. For example, if X is a cause of Y, and is not in part an effect of Y, r^2_{xy}, a coefficient of determination, gives the per cent of the variance of Y which is due to X. $1 - r^2_{xy}$, or k^2, gives the per cent of variance of Y which is due to other factors. For derivation and explanation, see:
   Tryon, R. C. "The Interpretation of the Coefficient of Correlation," Psychological Review, XXXVI (September, 1929), 419-45.

The first step of the path coefficient technique, after the intercorrelations between the different variables have been computed, is to set up a diagram in which the different variables are represented, for example, by numbered rectangles. These rectangles are then connected by arrows representing the assumed paths of influence and the directions of the influences between the variables. The arrows or paths are then designated by the symbols of their "path coefficients." The positions of the rectangles and the lengths of the arrows drawn in the diagram are without significance.

A path coefficient ($p_{1.2}$) is defined as the ratio of that part of the standard deviation of a variable which is due to another variable to the total standard deviation of the variable.[1] The path coefficient is equal to the coefficient of correlation between the two variables when only two variables are being considered. Wright has shown that when several variables are involved the coefficient of correlation between any two of them is equal to the path coefficient of the path directly connecting them plus the products of the path coefficients of the paths indirectly connecting them. For example, referring to Figure 1, r_{34} is equal to d the path of direct connection plus ab, the product of the paths a and b indirectly connecting 3 and 4. Paths leading to and through the dependent variable are not considered. This general theorem provides a basis for writing a series of simultaneous equations involving the intercorrelations and the path coefficients. Since the intercorrelations are known, solution of the equations gives the values of the path coefficients. The path coefficient values are then used in securing the coefficients of determination.

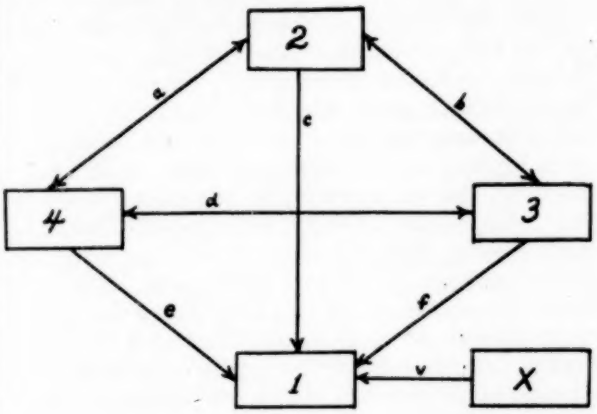


Figure 1
Variables and Paths of Influence

The coefficients of determination measuring the direct influences are secured by squaring the path coefficients of the paths leading directly from each of the independent variables to the dependent variable. The fundamental equation is:

$$d_{1.2} = p^2_{1.2}$$

The coefficients of determination measuring the joint or combined influences are obtained by taking twice the product of the two paths from the given pair of independent variables to the dependent variable times the

1. The interested reader should consult the first article by Wright for the derivation of the fundamental equations given here and for a more detailed explanation of the technique of path coefficients. The following discussion briefly indicates the origin of $d_{1.2} = p^2_{1.2}$ and $d_{1.\overline{23}} = 2p_{1.2}p_{1.3}r_{23}$. The variance of 1, σ_1^{-2}, is shown by Wright to be equal to $\sigma_2^2 + \sigma_3^2 + 2\sigma_2\sigma_3 r_{23}$ where 2 and 3 are two independent correlated variables in a system where 1 is the dependent variable. If each of these three terms is now divided by σ_1^2, the proportion of σ_1^2 (the variance of 1) which is due to 2 acting alone, 3 acting alone, and 2 and 3 acting jointly or in combination will be obtained. These ratios are the coefficients of determination and we may write: $d_{1.2} = \frac{\sigma_2^2}{\sigma_1^2}$; $d_{1.3} = \frac{\sigma_3^2}{\sigma_1^2}$; and $d_{1.23} = 2\frac{\sigma_2}{\sigma_1}\cdot\frac{\sigma_3}{\sigma_1}r_{23}$. Since the path coefficients $p_{1.2}$ and $p_{1.3}$ are equal respectively to $\frac{\sigma_2}{\sigma_1}$ and $\frac{\sigma_3}{\sigma_1}$, substitution yields the following equations: $d_{1.2} = p_{2.2}$, $d_{1.3} = p_{2.3}$, and $d_{1.\overline{23}} = 2p_{1.2}p_{1.3}r_{23}$.

coefficient of correlation between the two independent variables. The fundamental equation is:

$$d_{1.\overline{23}} = 2p_{1.2}p_{1.3}r_{23}$$

The sum of the coefficients of direct determination plus the sum of the coefficients of joint determination is equal to the total determination of the variance in the dependent variable which is due to the known independent variables. If all possible direct and joint influences are included the summation is equal to unity according to the fundamental equation:[1]

$$\Sigma d_{1.2} + \Sigma d_{1.\overline{23}} = 1$$

When all possible causal influences are not included the determination of the unknown causes may be represented by $d_{1.X}$ and the general equation becomes:

$$\Sigma d_{1.2} + \Sigma d_{1.\overline{23}} + d_{1.X} = 1$$

$\Sigma d_{1.2} + \Sigma d_{1.\overline{23}}$ then equals the total determination of the variance of the dependent variable by the known independent variables and is equal to the coefficient of multiple determination, or the square of the coefficient of multiple correlation between the dependent variable and the known independent variables.[2] $d_{1.X}$ is the coefficient of nondetermination and is equal to the square of the coefficient of multiple alienation. It may also be obtained by calculating the difference from unity of the other coefficients of determination.

The final step of the path coefficient technique involves the separation of the joint influences according to the weights indicated by the coefficients of direct determination. How this is accomplished is illustrated in the manipulation of the data of this study.

In Figure 1, the rectangles 1, 2, 3, and 4 represent the variables, problem solving ability, intelligence, computation ability, and reading ability as measured by the test scores. X represents unknown factors influencing problem solving ability. The arrows represent the assumed directions of the paths of influence. The small letters represent the path coefficients. Simple letters are used rather than the less convenient symbols $p_{1.2}$, $p_{1.3}$, and so on. The following assumptions were made in drawing Figure 1 and in writing the simultaneous equations from which the values of the path coefficients were obtained:

1. It is assumed that intelligence (2) contributes directly to variance in problem solving ability (1) through path c, and indirectly through reading ability (4) and computation ability (3) along paths ae and bf.

2. It is assumed that computation ability (3) contributes directly to variance in problem solving ability (1) through path f, and indirectly through intelligence (2) and reading ability (4) through paths bc and de. Another assumed indirect influence is along path bae through intelligence (2) and reading ability (4). The influence of computation ability on reading ability may, and probably is, insignificant, but it seems safest to include it.

3. It is assumed that reading ability (4) contributes directly to variance in problem solving ability (1) through path e, and indirectly through intelligence (2) and computation ability (3) along paths ac and df. Another assumed indirect influence is along path abf through intelligence (2) and computation ability (3).

4. It is assumed that reading ability and intelligence contribute to each other through path a.

5. It is assumed that computation ability and intelligence contribute to each other through path b.

6. It is assumed that unknown causes represented by X contribute to variance in problem solving ability through path v.

1. Where $\Sigma d_{1.2}$ includes $d_{1.2}$, $d_{1.3}$, and other coefficients of determination of this type, and $\Sigma d_{1.\overline{23}}$ includes $d_{1.\overline{23}}$ and other coefficients of the same type.
2. Wright has shown the relation between the techniques of path coefficients and the techniques of multiple correlation. Kelley has likewise demonstrated the relation. A path coefficient and a Beta regression coefficient are identical.
   Kelly, E. L. "The Relationship Between the Techniques of Partial Correlation and Path Coefficients," Journal of Educational Psychology, 20:119-24, February, 1929.

On the basis of these assumptions the writer uses the language of functional variation. For example, in stating that an independent variable contributes directly and indirectly to the variance of a dependent variable, he does so on the basis of a priori assumptions rather than on the basis of inferences drawn from the data.

It would be possible to assume other routes through which the influences could pass among the independent variables. Including these other routes adds to the complexity of the solution and, as Heilman points out, the computed path coefficient of each independent variable with the dependent variable, on which the conclusions of this study are based, would not be changed.[1]

In using coefficients of determination to measure the variance of a dependent variable which is due to an independent variable the assumption is made that the independent variable "contributes itself completely." That is to say, the independent variable is not in part an effect of the dependent variable, or in part an effect of more remote causes affected by the dependent variable. The influence of the dependent variable on an independent variable may be termed a reciprocal influence. The possible reciprocal influences of problem solving ability on intelligence as measured and computation ability as measured constitute a source of error to the conclusions of this study. Practice in problem solving undoubtedly affords training in computation. The Otis Test contains several items which are akin to verbal problems in arithmetic. Reciprocal influences from the dependent variable are not allowed for in the path coefficient techniques. We assume, essentially, that problem solving ability is intelligence plus other abilities. A possibly equally valid assumption is that intelligence is inclusive of problem solving ability.[2] or some portion of it. That is to say, intelligence is in part caused by or due to problem solving ability. Some support for the assumption that problem solving ability is intelligence plus other abilities is provided by the Spearman Two Factor Theory according to which problem solving ability would be analyzed as g + s, where "g" is considered as intelligence. If this theory is correct the assumption made in this study is justifiable to the extent that intelligence as measured is a cause rather than effect, or in part an effect, of problem solving ability. It should be noted in this connection that even if the dependent variable affects the independent variables the coefficients of determination are acceptable as minimum values. Hence, conclusions are not likely to overestimate the contributions ascribed to the variables intelligence, reading ability, and computational ability.[3] Reciprocal influences among the independent variables are taken care of in the path coefficient equations as written below.

The following simultaneous equations were written on the basis of the assumptions previously given. For example, r_{12} is equal to the path coefficient c, representing the path of direct connection, plus the product of the path coefficients a and e, representing one indirect connection, and plus the product of the path coefficients b and f representing another indirect connection. Other indirect connections adf and bde are not included. The simpler equations, neglecting some of these indirect connections, ultimately give the same results as would be obtained had all possible indirect connections been considered. It was stated previously that the computed path coefficient of each independent variable with the dependent variable, on which the conclusions of this study are based would not be changed.[4]

1. Heilman, op. cit., p. 58.
2. Tryon, op. cit., p. 423-24.
3. Support for this statement is found in the article by Tryon. A coefficient of determination in the general case where a common factor contributes unequally to the variance of two variables falls somewhere between the limits r^2 and r, where r represents the correlation between the two variables. Hence, r^2, which corresponds to the coefficients of determination used in this study, is the lower limit. Reciprocal influences may be thought of as due to a common factor which contributes unequally.
4. The equations given are identical with those used by Wright for a problem involving three independent variables in a system "in which each variable is connected with every other variable." The present writer has used different symbols and eliminated the parentheses in the equations for the path coefficients. These equations differ somewhat from those of Heilman, who includes more paths of indirect connection between the variables.

(Footnote continued)

$r_{24} = .9143 = a$
$r_{23} = .6447 = b$
$r_{34} = .5487 = d + ab$
$r_{12} = .6605 = c + ae + bf$
$r_{13} = .7806 = f + bc + de + bae$
$r_{14} = .5343 = e + ac + df + abf$

Solution of the equations given above resulted in the following values for the path coefficients:

$a = .9143$
$b = .6447$
$c = .5360$
$d = -.0407$
$e = -.2782$
$f = .5877$
$v^1 = .5796$

The values of the coefficients of determination were secured from the following equations based on the path coefficients just given.

$d_{1\cdot2} = c^2 = .2873$
$d_{1\cdot3} = f^2 = .3454$
$d_{1\cdot4} = e^2 = .0774$
$d_{1\cdot\overline{23}} = 2cfr_{23} = .4062$
$d_{1\cdot\overline{24}} = 2ecr_{24} = -.2727$
$d_{1\cdot\overline{34}} = 2efr_{34} = -.1795$
$d_{1\cdot X} = v^2 = .3359$
1.0000

Since $c = p_{1\cdot2}$, $f = p_{1\cdot3}$ the equation $d_{1\cdot\overline{23}} = 2cfr_{23}$ is the same as $d_{1\cdot\overline{23}} = 2p_{1\cdot2}p_{1\cdot3}r_{23}$. The same thing holds for $d_{1\cdot\overline{24}}$ and $d_{1\cdot\overline{34}}$.

The sum of $d_{1\cdot2}$, $d_{1\cdot3}$, $d_{1\cdot4}$, $d_{1\cdot\overline{23}}$, $d_{1\cdot\overline{24}}$, and $d_{1\cdot\overline{34}}$ represents the total determination of 1 due to factors 2, 3, and 4. The square root of this sum (.6641) is equal to .8149 and should equal $R_{1\cdot234}$. $R_{1\cdot234}$ when calculated from the coefficients corrected for attenuation with chronological age constant was found to be .8149. This indicates the accuracy of the computations since the value .8149 was secured from two entirely different series of calculations starting with the coefficients of correlation given in Table IV.

The coefficients of determination $d_{1\cdot2}$, $d_{1\cdot3}$, and $d_{1\cdot4}$ represent the direct influences of the causal factors intelligence, computation ability, and reading ability. The coefficients $d_{1\cdot\overline{23}}$, $d_{1\cdot\overline{24}}$, and $d_{1\cdot\overline{34}}$ represent the joint influences of the causal factors. The combined influence of intelligence and computation ability is represented by $d_{1\cdot\overline{23}}$. The combined influence of intelligence and reading ability is represented by $d_{1\cdot\overline{24}}$. The combined influence of computation ability and reading ability is represented by $d_{1\cdot\overline{34}}$. These two combined influences operate negatively, i.e., they tend to decrease the variance in problem solving ability.

When the combined, or joint, influences are separated on the basis of the relative weights of the direct influences and the results added to the magnitudes of the direct influences the variance of problem solving ability is attributed to the different independent variables as follows, in terms of per cents:[2]

1. 25.69 per cent of the variance in arithmetical problem solving ability is due to variation in intelligence as measured.

2. 42.05 per cent of the variance in arithmetical problem solving ability is due to variation in computation ability as measured.

Footnote 4 continued.
The values listed for path coefficients c, d, e, and f have been rounded off to four places. In calculating the coefficients of determination, they were used to five places in order to insure a closer check between the total of the coefficients of determination and the square of the multiple correlation coefficient obtained in the ordinary way. The coefficients of correlation used are those corrected for attenuation and with chronological age constant. More precisely the symbols should be $r_{12.5}$, $r_{13.5}$, and so on.

1. The value of v^2 was secured by difference from unity of the other coefficients of determination.
2. Separation of joint influences is illustrated as follows for $d_{1\cdot\overline{23}}$:

$$d_{1\cdot\overline{23}}\frac{d_{1\cdot2}}{d_{1\cdot2} + d_{1\cdot3}} = .4062\ \frac{.2873}{.2873 + .3454} = .18445$$

$$d_{1\cdot\overline{23}}\frac{d_{1\cdot3}}{d_{1\cdot2} + d_{1\cdot3}} = .4062\ \frac{.3454}{.2873 + .3454} = .22175$$

The value .18445 was added to $d_{1\cdot2}$ and the value .22175 added to $d_{1\cdot3}$. Separation of $d_{1\cdot\overline{24}}$ gave the values -.21483 and -.05787 which were combined with $d_{1\cdot2}$ and $d_{1\cdot4}$. Separation of $d_{1\cdot\overline{34}}$ gave the values -.14664 and -.03286 which were combined with $d_{1\cdot3}$ and $d_{1\cdot4}$.

3. -1.33 per cent of the variance in arithmetical problem solving ability is due to variation in reading ability as measured. That is to say, increasing variation in reading ability tends to decrease variation in arithmetical problem solving ability.

4. 33.59 per cent of the variance in arithmetical problem solving ability is due to the influence of other causes. These causes may include that which is peculiar to problem solving ability over and above intelligence, reading, and computation. They may include aspects of intelligence, reading, and computation unmeasured by our tests. They may include perseverance, interest, and other traits or characteristics which affect arithmetical problem solving ability.

The inferences may be drawn from this analysis of the variance in arithmetical problem solving ability that intelligence and computation ability are important factors in causing individual differences in problem solving ability. Since nothing can be done about intelligence, it would seem worthwhile to seek improvement in the solving of problems by providing instruction and drill in computation. It seems justifiable to infer that uniform instruction in computation, i. e., computation instruction in which all of the pupils participate is likely to increase individual differences in problem solving ability. The analysis does not indicate, however, the probable effect of remedial computation instruction administered to pupils who are inferior in problem solving ability. Remedial computation instruction administered to pupils whose problem solving achievement is inferior would in all likelihood tend to decrease individual differences in problem solving achievement. The achievement of the inferior pupils would be raised toward the central tendency of the whole group.

The analysis seems to indicate that general training in reading is likely to have a negligible, or possibly slight negative effect, on individual differences in problem solving achievement. A similar phenomenon has been observed in the case of supervised study. The inferior pupils are aided while the superior pupils are handicapped thus decreasing individual differences. It should be noted, however, that the reading test used measures a general type of reading achievement which possibly should not be expected to have a significant effect on problem solving achievement. It is possible, therefore, that variation in the type of reading ability which functions in the reading of arithmetical problems causes increased variance in arithmetical problem solving ability.[1] Training in the reading of arithmetical problems may result in increased individual differences where this training is administered to all the pupils. On the other hand, the administration of such training to the pupils who are inferior in arithmetical problem solving ability is likely to decrease individual differences in that the problem solving achievement of the inferior pupils is raised towards the central tendency.

The large amount of variance due to unknown causes should stimulate investigation of the possibility of other factors in problem solving ability. These unknown factors should be isolated and their effects measured, preferably by experiment.

It may be pointed out again that this study deals essentially with the variables intelligence scores, reading scores, computation scores, and problem scores. It is justified to refer to the traits as such only to the extent that the scores used are valid measures of the traits.

It should be noted that the variable chronological age was partialled out prior to the use of the technique of path coefficients. Hence, the conclusions stated above should be restricted to fifth-grade pupils of constant chronological age. Since the group of pupils used in this study may not be altogether typical of fifth-grade pupils in general, the conclusions stated above are not to be regarded as generalizations without qualification. Somewhat different results might be secured with a group which is more truly representative of fifth-grade pupils in general.

1. The findings of a recent study, in which partial correlations are given, support this conclusion. No functional relationship was reported between general reading ability and problem solving ability. A small but significant relationship was reported between arithmetical reading ability and problem solving ability.
B. A. Stevens. "Problem Solving in Arithmetic," Journal of Educational Research, 25, 253-60, April-May, 1932.

In conclusion the writer wishes to express his conviction that the path coefficient technique should be given wider use in the analysis of human abilities. Numerous problems suggest themselves. For example, the variance of ability to solve problems in chemistry might be analyzed to find out the relative importance of intelligence, arithmetical ability, knowledge of chemical formulae and equations, interest in chemistry, and so on. When the analysis has been completed experimentation should be used to determine the relative importance of controlled changes in the factors whose modification is possible.

AN EXPERIMENTAL ANALYSIS OF READING PERFORMANCE[1]

By Oscar F. Litterer
University of Minnesota

In general, three experimental methods have been employed in the analysis of reading performance: (1) the measurement of perception or visual apprehension with tachistoscopes; (2) the measurement of eye-movements by means of direct and indirect observation of the eyes in reading; and (3) paper and pencil tests constructed and standardized to measure individual and group performance in reading.

Few attempts have been made to discover the relationship that may be present among visual apprehension, eye-movements, and achievement on reading tests. The purpose of this study is to determine what degree of relationship exists among these measures of reading performance. To be more specific, eye-movement records and visual apprehension span will be validated as measures of reading performance with the paper and pencil tests as criteria. The relationship between span of visual apprehension and eye-movements will also be investigated; and since reliability of the techniques used affects directly the relationships studied, this factor will receive special attention.

The Minnesota eye-movement apparatus, as described by Tinker (10), was used to obtain the photographic eye-movement records. Two records of each subject were taken at one sitting. Fourteen lines of easy prose were read for the first record and fourteen lines of more difficult scientific prose for the second. Special precautions were employed to accustom the subject to the apparatus and to acquaint him with the general nature of the reading material. The first paragraphs of both types of prose were read by the subject before he was adjusted in the apparatus; the second paragraphs were read while the subject was in focus but no record was taken; and, finally, the third paragraphs were read for the records. In both the easy and the more difficult prose the three paragraphs formed connected discourse.

The subjects were instructed to read for an intelligent understanding of the subject matter, but not to study the selections. To check the comprehension of the material read, each subject was required to answer a set of questions on both trial and experimental paragraphs. Seventy-one readable records were secured for the easy prose and seventy-six for the more difficult prose. The fixation frequency, the perception time,[2] and the regression frequency per line were computed from the records. The relative span of apprehension was derived indirectly by dividing the number of words in the experimental paragraphs by the total fixation frequency.

The visual apprehension measurements were made with a modification of the Dodge Mirror Tachistoscope, described by Tinker (9), and with a simple exposure apparatus. In the former an exposure interval of 100 sigma was used. The simple exposure apparatus consisted of a stand constructed to hold large cards of printed material; a shield was used to conceal the material from the view of the subject and was lowered for the exposure. In this study, the exposure time was three seconds.

The reading material exposed in both apparatuses consisted of sentences. Those used with the Dodge Mirror Tachistoscope were composed of seven words and were approximately three and one-half inches in length; they were typed on smooth paper and pasted on strips of white cardboard. Sample: "This is my first day of work." The sentences employed with the simple exposure apparatus consisted of fifteen words and were hand-printed in three lines on large strips of white cardboard 11 by 28 inches in size, as for example: "The boy never stopped running nor looked behind him until he got to the house." With the Dodge Mirror Tachistoscope a preliminary series of 12 sentences which served to accustom the subject to the apparatus and method was given from four to seven

1. Part of a reading investigation done in the Psychological Laboratory, University of Minnesota, under the direction of Professor Miles A. Tinker.
2. Perception time is the pause duration multiplied by the pause frequency.

days before the experimental series of 70. With the other apparatus, an experimental series of 20 sentences immediately followed the preliminary series of five. The subjects were tested individually with the Dodge Mirror Tachistoscope, and in groups of 30 to 40 students each with the simple exposure apparatus. After the exposure was completed, the subject reproduced on paper as many of the words as possible. In both techniques, the score was computed on the basis of one point for a correct word reproduced in the proper order and one-half a point for a correct word in a changed order.

In silent reading, comprehension and speed are the essential variables to be measured. In this experiment, comprehension of reading was measured by the first part, paragraph comprehension, of the Iowa Silent Reading Test; and speed of reading was measured by two speed tests: the Chapman-Cook Speed of Reading Test, and the Minnesota Speed of Reading Test.

All subjects who participated in the investigation were sophomore psychology students. It was impossible to obtain measurements of each student on all parts of the experiment and, as a result, the number of cases range between 49 and 77 in the several comparisons. Table I gives the number of subjects secured on each of the measures, and the averages and standard deviations of the scores.

Since the results of this problem are mainly a function of intercorrelations, it was essential to secure measures of the reliability. For the eye-movement measures, the number of fixations and the perception time computed for the first seven lines of both selections were correlated against the same measures on the last seven lines of the same selection. This method has the advantage over the odd-even procedure in that each seven line units had the same number of words. Coefficients derived in this manner indicate the reliability for only half of the selections read. The reliability of the whole selections were obtained with the Spearman-Brown prophesy formula. These coefficients, as shown in Table II, are all .90 or above with the exception of that for the fixations on the more difficult prose (.775).

TABLE I

THE MEANS AND STANDARD DEVIATIONS OF THE SCORES

Measures	N	M	S.D.
Fixation Frequency on Easy Prose	71	6.00	1.17
Perception Time on Easy Prose	71	70.36	17.22
Regression Frequency on Easy Prose	71	.65	.50
Relative Span on Easy Prose	71	1.64	.31
Fixation Frequency on Scientific Prose	76	6.93	1.33
Perception Time on Scientific Prose	76	85.86	23.51
Regression Frequency on Scientific Prose	76	.96	.59
Relative Span on Scientific Prose	76	1.30	.27
Visual Appre. Scores - 100 sigma exposure	77	3.43	.90
Visual Appre. Scores - three sec. exposure	70	12.14	1.85
Chapman-Cook Speed of Reading Test	74	17.71	3.65
Minnesota Speed of Reading Test	49	26.83	5.15
Iowa Silent Reading Test	72	57.15	14.49

TABLE II

RELIABILITY COEFFICIENTS

Measures	Raw r	P.E.	S - B Prophesy Formula	P.E.
Fixation Frequency on Easy Prose	.854	.02	.921	.01
Perception Time on Easy Prose	.834	.02	.909	.01
Fixation Frequency on Scientific Prose	.633	.04	.775	.03
Perception Time on Scientific Prose	.839	.02	.912	.01
Visual Appre. Scores - 100 sigma Exposure	.937	.01	.967	.00
Visual Appre. Scores - three sec. Exposure	.837	.02	.911	.01
Chapman-Cook Speed of Reading Test	.780	.03	.755[1]	.03
Minnesota Speed of Reading Test	.860	.02	.798[1]	.03
Iowa Silent Reading Test	.773	.03	.871	.02

1. These coefficients were taken from previous studies (7, 1) made on comparable groups and corrected for the difference in the standard deviations by Kelley's (5) formula for adjusting reliability coefficients.

The reliabilities of the visual apprehension scores were obtained by correlating the scores of the first half of each test with those of the second half. The coefficients of .967 and .911 imply that the visual apprehension span was measured with much consistency either by the simple exposure apparatus or by the Dodge Mirror Tachistoscope.

In the paragraph comprehension test, the items in the first and second half are not comparable and, therefore, the odd-even method was used to obtain the reliability. The reliability coefficient of this measure proved to be .871 which reveals nearly as much consistency as the eye-movement measures and the visual apprehension scores.

One form of the Chapman-Cook Speed of Reading Test and the Minnesota Speed of Reading Test were given, hence the impossibility of securing a reliability coefficient. An average coefficient for the former was obtained from Tinker and Paterson's (7) investigations, and for the latter from Eurich's (1) study. They were corrected for the difference in the standard deviations by Kelley's (5) formula for adjusting reliability coefficients. The coefficients for these tests are .755 and .798 which are somewhat lower than those obtained on the other measures. They, nevertheless, possess sufficient consistency for the analysis of group performance.

Each eye-movement measure for the easy prose was correlated with the corresponding record for scientific prose to determine the agreement of the measures for the different types of reading matter. The four measures, consisting of the fixation frequency, the perception time, the relative span, and the regression frequency, as tabulated in Table III, have correlations which show a fair degree of relationship. The raw correlations range from .674 for regression frequency to .832 for perception time. The fixation frequency and the relative span correlated to an equal degree (.77) because the same variable is involved in both measures, namely, the number of fixations. The raw correlations, with the exception of the one for the regression frequency, corrected for attenuation, are approximately .91 for the three eye-movement measures. Fixation frequency and perception time, according to the corrected correlations, measure eye-movements with the same degree of consistency when the errors of measurement are eliminated. Tinker (8) has shown that the greater the change in type of reading material and the resulting reading attitude, the smaller is the correlation between the eye-movement measures of two different selections. This interpretation suggests that the two selections used in this study were read by the subjects with about the same attitude.

TABLE III

INTERCORRELATIONS ON THE SIMILAR MEASURES

Measures	Raw r	P.E.	Atten. r	P.E.
Fixation Frequency on Easy Prose vs. Fixation Frequency on Scientific Prose	.775	.03	.918	.01
Perception Time on Easy Prose vs. Perception Time on Scientific Prose	.832	.02	.914	.01
Regression Frequency on Easy Prose vs. Regression Frequency on Scientific Prose	.674	.04	---	---
Relative Span on Easy Prose vs. Relative Span on Scientific Prose	.773	.03	.915	.01
Visual Appre. Scores - 100 sigma exposure vs. Visual Appre. Scores - three sec. exposure	.570	.05	.607	.05
Chapman-Cook Speed of Reading Test vs. Minnesota Speed of Reading Test	.620	.05	.798	.03
Chapman-Cook Speed of Reading Test vs. Iowa Silent Reading Test	.565	.05	.697	.04
Minnesota Speed of Reading Test vs. Iowa Silent Reading Test	.540	.06	.648	.05

The correlation of .570 (corrected for attenuation, .607) reveals only a moderate relationship between the two visual apprehension tests. The short exposure of 100 sigma excluded all eye-movements and made it possible to secure a visual apprehension span of one fixation. The three second exposure, on the other hand, yielded a span involving several fixations. It is evident that the two measures of apprehension are not interchangeable.

The Minnesota Speed of Reading Test measures rate of reading with more difficult content than the Chapman-Cook Test. The former was constructed to measure the rate at which college students read content comparable to their general reading. The Chapman-Cook Test was designed to measure the reading rate of elementary pupils. This test, however, has been widely used as a measure of the reading rate on the high school and college level. On the college level the comprehension errors in the Chapman-Cook are practically zero. The correlation between the scores of the two speed tests is .620 which is the same as that

found by Eurich (1). The coefficient for the present data, corrected for attenuation is .798. Paragraph comprehension, the first part of the Iowa Silent Reading Examination, shows a low but significant relationship to the two speed of reading tests. The correlation is .565 with the Chapman-Cook and .540 with the Minnesota Speed Test. The coefficients corrected for attenuation are .697 and .648 respectively.

Further intercorrelations are given in Table IV. The eye-movement measures show a low but rather uniform relationship to the visual apprehension scores.

is not very diagnostic of apprehension in the more normal situation which involves continuous reading of related material.

A series of experiments have been performed to determine the importance of perception in reading. Gates (2) has interpreted his data --secured from elementary school subjects-- to mean that perception is specialized rather than general. Mueller, cited by Starch, (6), with high school students as subjects, obtained correlations of .64, .73 and .70 respectively when the apprehension span for unrelated words was correlated against speed, comprehension, and speed plus comprehension

TABLE IV

INTERCORRELATIONS OF VISUAL APPREHENSION, EYE-MOVEMENTS AND READING TEST SCORES. IN EACH CASE THE COEFFICIENT, CORRECTED FOR ATTENUATION, FOLLOWS THE RAW COEFFICIENT[1]

	Visual Appre. 100σ Exposure		Visual Appre. 3 sec. Exposure		Chapman-Cook Speed of Reading Test		Minnesota Speed of Reading Test		Paragraph Comprehension of the Iowa Silent Reading	
Measures	Raw r	Atten. r	Raw r	Atten. r	Raw r	Atten. r	Raw r	Atten. r	Raw r	Atten. r
1. Fixation Frequency on Easy Prose	-.437	-.463	-.313	-.342	-.433	-.519	-.327	-.381	-.476	-.531
2. Perception Time on Easy Prose	-.343	-.366	-.431	-.474	-.496	-.599	-.285	-.334	-.413	-.464
3. Relative Span on Easy Prose	.463	.490	.328	.358	---	---	---	---	---	---
4. Fixation Frequency on Scientific Prose	-.383	-.442	-.295	-.351	-.513	-.671	-.326	-.414	-.450	-.548
5. Perception Time on Scientific Prose	-.326	-.347	-.342	-.375	-.460	-.554	-.270	-.316	-.393	-.441
6. Relative Span on Scientific Prose	.356	.441	.293	.348	---	---	---	---	---	---
7. Visual Appre. - 100 σ Exposure					.513	.600	.549	.625	.512	.558
8. Visual Appre. - 3 sec. Exposure					.423	.510	.405	.475	.421	.473

1. The average probable error is ± .06 with range from ± .04 to ± .08.

The average correlation is .384 with a range from .326 to .463[1] between the eye-movement measures and the visual apprehension scores secured with a 100 sigma exposure. The average corrected coefficient is .424 with a range from .347 to .490. A similar average and range was obtained between the eye-movement measures and visual apprehension scores secured with a three second exposure. The average correlation is .333 with a range from .313 to .431, and an average of .374 with a range from .342 to .474 when corrected for attenuation. Apparently, apprehension of material exposed for 100 sigma or three seconds

in reading. When span of related words was substituted for unrelated words, the correlations were .70, .59 and .69 respectively with the same measures. In the present investigation college students were employed as subjects, and a somewhat lesser relationship was found between apprehension scores and reading test scores. The Chapman-Cook correlated .513 with apprehension scores secured with a 100 sigma exposure and .423 with a three second exposure. The corrected coefficients are .600 and .510 respectively. The Minnesota Speed Test shows practically the same relation to the apprehension scores.

1. It seemed more adequate to ignore signs in averaging these coefficients.

The correlations are .549 and .405 respectively for the 100 sigma and three second exposure; and the corrected coefficients are .625 and .475 between the same measures. Similar coefficients were obtained when apprehension scores were compared with comprehension in reading. The correlations between paragraph comprehension and apprehension scores secured with 100 sigma and three second exposure are .512 and .421 respectively. These correlations corrected for attenuation are .558 and .473. These data suggest that the span of apprehension plays a less important part in the reading of adult subjects than with high school and elementary school pupils.

It has been assumed that fixation frequency and perception time are good measures of speed and comprehension in reading. Gray (3) found that the slowest reader made longer and more fixations than the most rapid reader. In this investigation, however, the correlations reveal only a moderate relationship. The fixation frequency on the easy and scientific prose correlated -.433 and -.513, and when corrected for attenuation -.519 and -.617 with the Chapman-Cook Speed of Reading Test. Mueller, cited by Starch, (6) found a similar relationship,[1] namely, a coefficient of -.54 between the number of fixations and rate of reading. The Minnesota Speed of Reading Test shows a slightly smaller correlation with fixation frequency. The coefficients are -.327 and -.326, and corrected for attenuation are -.381 and -.414 respectively with the easy and more difficult prose.

Perception time, according to Tinker (8), is the most valid photographic measure of rate in reading when all measures are on the same reading material. In this experiment, both perception time and fixation frequency, as measured in reading material different than that in the reading tests, show approximately the same correlation with speed of reading. The coefficients are -.285 and -.270, and when corrected for attenuation -.334 and -.316 between the easy and more difficult prose respectively, and the Minnesota Speed of Reading Test. The Chapman-Cook Speed of Reading Test correlated somewhat higher with the same photographic measures. These correlations of -.496 and -.460 when corrected for attenuation are -.599 and -.554.

Paragraph comprehension, as measured by the Iowa Silent Reading Examination, shows a fair relationship to the two eye-movement measures. The fixation frequency on the easy and more difficult prose correlated -.476 and .450, and corrected for attenuation -.531 and -.548 with paragraph comprehension. Mueller, cited by Starch, (6) obtained between comprehension and fixation frequency a correlation of -.49 which is slightly higher. Perception time in this study correlated with the same reading test -.413 on the easy and -.393 on the more difficult prose. The coefficients corrected for attenuation are -.464 and -.441. Eye-movement records, therefore, correlated about as well with comprehension as with rate in reading.

Although these intercorrelations when corrected for attenuation are not raised much (average raw r is -.403; average corrected r is -.481), as can be observed by a comparison of the raw and corrected coefficients in Table IV, the true relationships have some significance. An average corrected coefficient of -.473 with a range from -.316 to -.671 was obtained between the eye-movements and rate of reading. A similar average coefficient of -.496 with a range from -.441 to -.548 was secured between eye-movements and comprehension. One can tentatively conclude that eye-movements on easy and scientific prose, and reading tests tend to measure somewhat similar reading performance.

SUMMARY AND CONCLUSION

1. In this investigation, the attempt was made to determine the approximately true relationship between visual apprehension, eye-movements, and achievement on reading tests.

2. Reliability of the measuring devices was given special attention. The measures yielded a reliability sufficiently high for group comparisons. The eye-movement measures and the visual apprehension scores, in particular, showed a very high degree of internal consistency.

1. Starch lists the correlations as positive but writes in the discussion that the fixation frequency per line is probably inversely proportional to the amount apprehended at each fixation. They are cited as negative in this report to make them comparable to our correlations.

3. Each eye-movement measure showed a high correlation between easy and scientific prose, indicating that the two paragraphs were not read with markedly different attitudes.

4. Although the visual apprehension scores secured by exposure intervals of 100 sigma and three seconds show a fair degree of agreement with each other (.607), they are not interchangeable.

5. The scores on the three reading tests: the Chapman-Cook Speed of Reading Test, the Minnesota Speed of Reading Test, and paragraph comprehension intercorrelated to an appreciable degree (.648 to .798).

6. Eye-movements showed only a moderate relationship with the visual apprehension scores (-.342 to -.474).

7. The correlations, corrected for attenuation, between visual apprehension scores and the reading test scores reveal a fair degree of relationship (.473 to .625).

8. Fixation frequency and perception time in reading is correlated only moderately well with rate and comprehension.

9. In general, the data justify the conclusion that there is a fairly significant but not high relationship among visual apprehension scores, eye-movements, and achievement test scores in reading.

BIBLIOGRAPHY

1. Eurich, A. C. The Reading Abilities of College Students, University of Minn. Press, Minneapolis, 1931, pp. 56, 57.
2. Gates, A. I. "A Study of the Role of Visual Perception, Intelligence, and Certain Associative Processes in Reading and Spelling," Journal of Educational Psychology, (1926), 17, 433-445.
3. Gray, C. T. "Types of Reading Ability as Exhibited Through Tests and Laboratory Experiments," Supplementary Educational Monograph, No. 5 (1917).
4. Hoffman, J. "Experimentelle-psychologische Untersuchungen ueber Leseleistungen von Schulkindern," Archives f. d. gestalte Psychologie, 1927, 58, 325-388.
5. Holzinger, K. J. Statistical Methods for Students in Education, Ginn and Company, Boston, Massachusetts, 1928, 253-254.
6. Starch, D. Educational Psychology, Macmillan Company, New York, 1919, 281-296.
7. Tinker, M. A. & Paterson, D. G. "Studies of Typographical Factors Influencing Speed of Reading," Journal of Applied Psychology, Vol. XV, No. 5 (Oct. 1931) 471-479.
8. Tinker, M. A. "Photographic Measure of Reading Ability," Journal of Educational Psychology, (March, 1929); 184-191.
9. Tinker, M. A. "A Noiseless Exposure Apparatus," American Journal of Psychology, Vol. XLIII (Oct. 1931), 640-642.
10. Tinker, M. A. "Apparatus for Recording Eye-Movements," American Journal of Psychology, Vol. XLIII, (January,1931) 115-118.
11. Tinker, M. A. "Visual Apprehension and Perception in Reading," Psychology Bulletin, Vol. XXVI, No. 4 (April, 1929). 223-240.

A STUDY OF THE RELATIVE AMOUNTS OF TRANSFER RESULTING FROM THREE METHODS OF STUDY

By John B. Cluley
University of Wisconsin

The purpose of the investigation reported herein is to determine the relative amounts of transfer resulting from three methods of studying rectangles and triangles.

The practical utility of most school studies has been shown quite conclusively to be rather limited and, in most instances, to follow the law of diminishing returns. E. L. Thorndike sets forth the premise that even a small amount of transfer effect may equal in importance, after a certain point, the direct utilities of the school studies.[1] In two very extended investigations he shows that the differences in transfer effect resulting from different school studies is small. If transfer, however, depends, as C. H. Judd says it does, upon the extent to which the information is generalized, then it is not unreasonable to believe that the method of study may have quite as much to do with the amount of transfer as does the subject matter studied. Several recent investigations seem to indicate this possibility.[2]

Procedure and Materials

The equivalent-groups method of experimentation was used in the study. Three groups of pupils, 95 cases in each group, were obtained from the sixth and seventh grades of four cities in southern Wisconsin. The pupils were matched on the basis of mental age and intelligence quotient, as measured by the Detroit Alpha Intelligence Test, Form M.

A fixed time-schedule was followed in each room to avoid irregularity and interruptions of work. With the cooperation of the teacher in charge of the room, the investigator made out a schedule before any of the work was done. The schedule was followed without variation in every room where work was carried on. A copy of a schedule used in one of the rooms is given for illustration.

Time-Schedule Used in a Stoughton, Wisconsin, Sixth Grade.

Monday,	Oct. 19	9:00-9:45	Detroit Alpha Int. Test
Tuesday,	Oct. 20	9:00-9:20	Arithmetic Pretest
Monday,	Oct. 26	9:00-9:30	Study of rectangles
Tuesday,	Oct. 27	9:00-9:30	Study of triangles
Wednesday,	Oct. 28	9:00-9:20	Arith. Final Test
Wednesday,	Nov. 4	9:00-9:20	Arith. Retention Test

The subject matter of the investigation dealt with rectangles and triangles. Three phases of the material were considered: (1) general information; (2) recognition of the forms; and (3) computation of areas. The same subject matter content was presented by three different methods of instruction. The instruction was made uniform for all pupils by using mimeographed materials. The three methods, usually spoken of in the tables as methods A, B and C, are described in detail.

Method A was planned to represent a very formal presentation of the information on rectangles and triangles. The formula for finding the area of each figure was given and the children were instructed to find the areas by means of it. Pupils were given the terms base and altitude without any objective helps as to their meaning. No suggestions were made regarding the fact that common figures and spaces are made up of rectangles and triangles. Verbal problems dealing with abstract situations were used as practice material in this method of study.

1. Thorndike, E. L. "Instruments for Measuring the Disciplinary Values of Studies," Journal of Educational Research, V, 1922, 269-279.
2. Woodrow, Herbert. "The Effect of Type of Training on Transference," Journal of Educational Psychology, XVIII, 1927, 159-172.
   Overman, J. R. An Experimental Study of Certain Factors Affecting Transfer of Training in Arithmetic (Warwick and York, 1931).
   Meredith, C. P. "Consciousness of Method as a Means of Transfer of Training," Forum of Education, V, 1927.
   Coxe, W. W. "The Influence of Latin on Spelling of English Words," Journal of Educational Research Monograph, No. 7, 1925, p. 121.

Method B is based upon an objective presentation of the information on rectangles and triangles. The area of a rectangle is shown to be the number of squares formed in the figure when lines are drawn, one linear unit apart, both ways across the rectangle. A triangle is shown, by a pictorial method, to be one half of a rectangle whose dimensions are the same as those of the triangle. The problems used for practice in computing areas are verbal but employ rectangles and triangles in their natural settings in the child's environment.

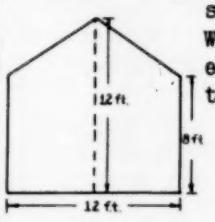
Method C is a variation of Method B. The same method of presenting information is employed. The same instruction on finding the areas is given as in B. The practice work is, however, of two kinds: first, the pupils are given an opportunity to recognize rectangles and triangles in larger figures; second, pupils are requested to find areas of rectangles and triangles which are shown to be part of a larger figure. In some cases the words rectangle and triangle are not used in connection with the picture. The problems are not of the verbal type but refer to the common figures and spaces in the child's environment as do the problems in method B.

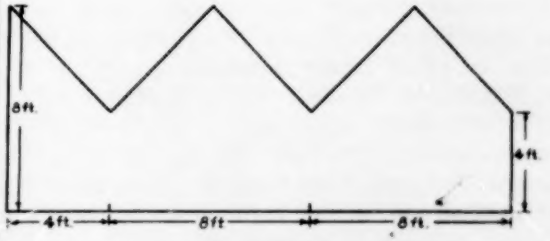
A pretest of four parts was given to all pupils before the study material employed in this investigation was presented. This test aims only to determine the amount of information concerning rectangles and triangles possessed by the pupils and their ability to handle the forms in problems. The test does not aim to measure general mathematical ability. Part I of the test calls for general information about rectangles and triangles. It is almost entirely of the completion type. Most of the blanks to be filled do not permit of any variation in response. Part II of the test is composed of verbal problems on rectangles and triangles. They are of the same type as those used in practice in method A. Part III of the test is made up of verbal problems which deal with real situations. They are the same as those used in the method B practice material. Part IV of the test includes a series of picture problems which correspond with the practice material of method C. An illustration of each type of problem is given below:

Method A problem: "2. A rectangle is 18 feet long and 14 feet wide. Find the area of the rectangle. ..... 5. Find the area of a triangle that has a base of 16½ rods and an altitude of 12 rods."

Method B problems: "2. The dimensions of a window are 3 feet and 6 feet. What is the area of the window? ..... 5. An odd shaped piece of land is divided into two triangles. Each triangle has a base of 15 rods. One has an altitude of 12 rods, the other an altitude of 8 rods. What is the area of the piece of land?"

Method C problems: "2. The figure represents the end of a garage. What is the area of the largest rectangle you can draw in the figure?





Find the area of this figure."

The scoring of the test requires some explanation. Problems of this sort are usually considered correct only when the correct answer is obtained. A correct answer is, however, no guarantee that the pupil has used the correct reasoning in obtaining it. Some of the test papers show that correct answers were obtained by very faulty operations. Since this investigation does not have to do with general mathematical ability, it appeared to the writer that the scores made on the test should not be influenced by the lack of the mastery of the fundamental operations or by mistakes in performance due to nervousness or careless habits. The scores should constitute an index of the pupil's ability to recognize the form and his knowledge of how to compute the areas. For this reason, Parts II, III and IV of the test were scored on the

basis of one point for each time the pupil used the correct operation in the problems.

Two study periods were given the pupils. During the first period all pupils studied rectangles and in the second they all studied triangles. Every child in the room, whether in a matched group or not, took part in the activity going on at the time. The pupils were instructed to gain as much information as it was possible for them to do during the thirty minutes study period. The pupils were told that it was not necessary to finish the three pages of study material. The person in charge of the study periods gave no help to the pupils other than to make clear what was printed on the papers in their possession.

A final test followed the study periods. The same test as that used in the pretest was given as the final test. In treating the data, it is assumed that gains in scores made on the different parts of the final test, over the pretest, may in some instances be attributed to direct instruction and practice while in other instances the gains are attributed to transfer effect, depending upon the method of study used by the pupils making the gains. The test was given at a later date as a retention test to determine the effectiveness of the three methods of study with respect to retention.

The reliability of the test was determined by two different methods: the self-correlation method and the chance-halves method. In the first method the scores made by the same pupils at two different times, under approximately the same conditions, were used as the basis of computation. The scores on the final test and the retention test were used for this purpose. The results are given in Table I.

TABLE I

COEFFICIENTS OF CORRELATION BETWEEN SCORES MADE ON THE FINAL RETENTION TESTS BY SEVEN GROUPS OF PUPILS

Grade	No. of cases	r	P.E.$_r$
Sixth grade	36	.6537	.064
Sixth grade	74	.8371	.024
Sixth and seventh	45	.8349	.030
Sixth and seventh	60	.8742	.021
Seventh	63	.8973	.019
Seventh	60	.9390	.010
Sixth and seventh (from the above groups)	271	.9395	.004

The reliability of the test was also determined for two groups of pupils (first and sixth groups of Table I) by means of the chance-halves method. The scores on the odd and even numbered points of the final test were used in this method. A coefficient of reliability of .8923 $\pm$ .023 was obtained for the sixth grade group of 36 pupils. The scores for the seventh grade group of 60 pupils yielded a coefficient of reliability of .9549 $\pm$ .008.

Results

In order to overcome such inequalities as may exist in the unit of measurement employed in this investigation the results for the investigation have been presented in five different ways: (1) the raw scores (Table II); (2) standard scores, transmuted on the basis of the pretest (Table III); (3) standard scores, transmuted on the basis of the final scores for the practice groups (Table IV); (4) the per cent of gain with the pretest as the base (Table V) and (5) the per cent of gain with the possible gain as the base (Table VI).

The average gains on the final test over the pretest are statistically significant for all study groups on each part of the test. The 27 critical ratios given in Tables II, III and IV range from 3.7 to 8.4, the three methods of treating the scores giving practically the same results. The data given in Tables V and VI also appear to represent real gains. While the gains are statistically significant, in the absence of positive evidence to the contrary, there is the possibility that they may be due to practice effect from a second application of the test employed in the investigation.

In the absence of direct evidence as to the amount of practice effect, the gains made on the final test over the pretest were compared with the gains made on the retention test over the final test (Table VII). The gains made on the retention test over the final test are not due to instruction. Since the differences between the average gains on the final test over the pretest and the average gains on the retention test over the final test are not, with one exception, statistically significant, it would seem that the gains on the final test over the pretest may have been due to practice effect from the second application of the test rather

than to instruction. While this may be true the addition of a constant amount (practice effect) to all groups in no manner invalidates the comparison of the results from one group to another insofar as these differences are statistically significant.

A comparison of the amounts of transfer secured from the three methods is presented in Table VIII. While none of the differences are statistically significant, Method B proves to be superior to Method A in seven of the nine comparisons, two of the critical ratios being 2.2. In the two comparisons which do not favor Method B, the critical ratios are practically zero (.029 and .032). Method B is slightly superior to Method C and Method C is somewhat superior to Method A. The small differences between the several methods would seem to indicate that two thirty minute study periods were insufficient, the small gains from instruction being lost, for comparative purposes, in the practice effect, resulting from two applications of the test.

A more detailed analysis of the data brings out several points of interest. In the first place it will be noted that none of the methods of study employed in this investigation seem to be effective in teaching pupils to solve problems of types III and IV. In general Method B is the most effective for all types of problems. On the final test the transfer gains appear to be greater than the instructional gains, while on the retention test the practice groups exceed the transfer groups regardless of the method employed in study. Whether this difference in gains may be the result of chance factors or whether it may be due to the short study period used in the investigation or some other cause the writer is unable to say.

CONCLUSIONS

1. Of the three methods of instruction studied in this investigation, Method B appears to be superior to Methods A and C in teaching pupils to solve all types of problems used in the investigation. The transfer effect for Method B appears to be greatest.

2. None of the methods of study used in the investigation appear to be particularly effective in teaching pupils to solve problems of types III and IV.

3. The time allotted to study in the investigation was probably not long enough to produce instructional gains and transfer gains sufficiently large, considering the large gains from the practice effect of two applications of two tests, to be considered statistically significant.

TABLE II

TABLE SHOWING TOTAL SCORES, AVERAGE SCORES, GAINS IN AVERAGE SCORE OVER PREVIOUS TEST, AND STATISTICAL SIGNIFICANCE OF GAINS
(Critical Ratio computed on basis of probable error)

Test		Pretest			Final			Retention		
Part of test		II	III	IV	II	III	IV	II	III	IV
Method A	Total score	541	352	63	807	476	103	969	548	144
	Average score	5.69	3.7	.66	8.49	5.01	1.08	10.2	5.77	1.51
	Average gain over previous test				2.80	1.31	.42	1.53	.76	.43
	Statistical significance of gain (C.R.)				5.8	4.04	4.7	2.9	1.8	2.0
Method B	Total score	415	279	66	734	403	137	798	500	169
	Average score	4.37	2.94	.69	7.73	4.24	1.44	8.4	5.26	1.78
	Average gain over previous test				3.36	1.30	.75	.67	1.02	.34
	Statistical significance of gain (C.R.)				8.4	4.6	4.8	2.2	2.3	1.8
Method C	Total score	546	293	57	795	450	111	910	513	153
	Average score	5.77	3.08	.60	8.37	4.74	1.16	9.58	5.40	1.61
	Average gain over previous test				2.60	1.66	.56	1.21	.66	.45
	Statistical significance of gain				5.1	5.2	4.1	1.7	2.3	2.3

TABLE III

COMPARABLE SCORES, TRANSMUTED ON THE BASIS OF SCORES MADE ON THE PARTS OF THE PRETEST, SHOWING TOTAL SCORES, AVERAGE SCORES, GAINS IN AVERAGE SCORE OVER PREVIOUS TEST, AND STATISTICAL SIGNIFICANCE OF GAINS

Test		Pretest			Final			Retention		
Part of test		II	III	IV	II	III	IV	II	III	IV
Method A	Total score	541	547.2	254.6	807	749.6	408.5	969	866.4	565.3
	Average score	5.69	5.76	2.68	8.49	7.89	4.30	10.2	9.12	5.95
	Gain in av. score over previous test				2.80	2.13	1.62	1.71	1.23	1.65
	Statistical significance of gain				6.3	4.8	3.7	3.9	2.8	3.7
Method B	Total score	415	430.4	266	734	630.8	539.6	798	787.6	663
	Average score	4.37	4.53	2.80	7.73	6.64	5.68	8.40	8.29	6.99
	Gain in av. score over previous test				3.36	2.11	2.88	.67	1.65	1.31
	Statistical significance of gain				7.6	4.8	6.5	1.5	3.7	2.9
Method C	Total score	546	452.2	232.8	795	707.8	438	910	809.4	602.3
	Average score	5.77	4.76	2.45	8.37	7.45	4.61	9.58	8.52	6.34
	Gain in av. score over previous test				2.60	2.69	2.16	1.21	1.07	1.73
	Statistical significance of gain				5.9	6.1	4.8	2.7	2.4	3.9

*Note.--The total and average scores of Table II were obtained by transmuting the scores of Table I into comparable scores on the basis of the scores made on the pretest by all the pupils.

TABLE IV

COMPARABLE SCORES, TRANSMUTED ON THE BASIS OF THE SCORES MADE ON PARTS II, III AND IV BY GROUPS A, B AND C, RESPECTIVELY, SHOWING TOTAL SCORES, AVERAGE SCORES, GAIN IN AVERAGE SCORE OVER PREVIOUS TEST, AND THE STATISTICAL SIGNIFICANCE OF THE GAIN.

Test		Pretest			Final			Retention		
Part of test		II	III	IV	II	III	IV	II	III	IV
Method A	Total score	541	626.6	296.8	807	860.7	562.4	969	998.5	736.3
	Average score	5.69	6.59	3.93	8.49	9.06	5.92	10.2	10.49	7.73
	Gain in av. score over previous test				2.80	2.47	1.99	1.71	1.43	1.83
	Statistical significance of gain				5.3	4.7	3.8	2.9	2.7	3.5
Method B	Total score	415	489.3	385.7	734	722	706.8	798	905.4	852.2
	Average score	4.37	5.15	4.06	7.73	7.60	7.44	8.40	9.53	8.97
	Gain in av. score over previous test				3.36	2.45	3.38	.67	1.93	1.53
	Statistical significance of gain				6.3	4.6	6.4	1.3	3.6	2.9
Method C	Total score	546	514	347.7	795	812.3	587	910	930	779
	Average score	5.77	5.41	3.66	8.37	8.55	6.18	9.58	9.79	8.20
	Gain in av. score over previous test				2.60	3.14	2.52	1.21	1.24	2.02
	Statistical significance of gain				4.9	5.9	4.8	2.3	2.3	3.8

*Note.--The total and average scores of Table III were obtained by transmuting the scores of Table I into comparable scores on the basis of the scores made on Parts II, III and IV of the final test by Group A, Group B and Group C, respectively.

TABLE V

A TABLE SHOWING THE GAIN IN AVERAGE FINAL TEST SCORE OVER AVERAGE PRETEST SCORE, GAIN IN AVERAGE RETENTION TEST SCORE OVER AVERAGE FINAL TEST SCORE, AND THE PER CENT THE GAIN IS OF THE POSSIBLE GAIN

	Method A			Method B			Method C		
Part of test	II	III	IV	II	III	IV	II	III	IV
Av. final test score	8.49	5.01	1.08	7.73	4.24	1.44	8.37	4.74	1.16
Av. pretest score	5.69	3.70	.66	4.37	2.94	.69	5.77	3.08	.60
Average gain	2.80	1.31	.42	3.36	1.30	.75	2.60	1.66	.56
Possible gain	23.31	13.3	28.34	24.63	14.06	28.31	23.23	13.92	28.40
Per cent of possible gain	12.00%	9.85%	1.48%	13.63%	9.24%	2.65%	11.19%	11.92%	1.97%
Av. retention test score	10.2	5.77	1.51	8.40	5.26	1.78	9.58	5.40	1.61
Av. final test score	8.49	5.01	1.08	7.73	4.24	1.44	8.37	4.74	1.16
Average gain	1.71	.76	.43	.67	1.02	.34	1.21	.66	.45
Possible gain	20.51	11.99	27.92	21.27	12.76	27.56	20.63	12.26	27.84
Per cent possible gain	8.34%	6.34%	1.54%	3.14%	7.99%	1.23%	5.86%	5.38%	1.62%

TABLE VI

A TABLE SHOWING AVERAGE SCORES, AVERAGE GAINS, AND THE PER CENT THE AVERAGE GAIN IS OF THE AVERAGE PRETEST SCORE ON EACH PART OF THE TEST FOR EACH METHOD OF STUDY

	Method A			Method B			Method C		
Part of test	II	III	IV	II	III	IV	II	III	IV
Av. score on final test	8.49	5.01	1.08	7.73	4.24	1.44	8.37	4.74	1.16
Av. score on pretest	5.69	3.70	.66	4.37	2.94	.69	5.77	3.08	.60
Average gain	2.80	1.31	.42	3.36	1.30	.75	2.60	1.66	.56
Per cent av. gain is of average pretest score	49.2%	35.2%	63.6%	76.9%	45.2%	108%	47.9%	53.9%	94.9%
Av. score on retention test	10.2	5.77	1.51	8.40	5.26	1.78	9.58	5.40	1.61
Av. score on final test	8.49	5.01	1.08	7.73	4.24	1.44	8.37	4.74	1.16
Average gain	1.71	.76	.43	.67	1.02	.34	1.21	.66	.45
Per cent average gain is of average pretest score	30.05%	20.54%	65.15%	15.3%	34.7%	49.9%	21.97%	21.43%	75.00%

TABLE VII

COMPARISON OF THE AVERAGE GAINS ON FINAL TEST OVER PRETEST WITH THE AVERAGE GAINS ON RETENTION TEST OVEH FINAL TEST

	Part of test	II	III	IV
Method A	Av. gain, final over pretest	2.80 ± .460	1.31 ± .307	.42 ± .092
	Av. gain, retention over final	1.71 ± .592	.76 ± .346	.43 ± .140
	Difference in gain	1.09	.55	-.01
	P. E. of difference	.75	.463	.168
	Critical Ratio	1.45	1.19	.06
	Part of test	II	III	IV
Method B	Av. gain, final over pretest	3.36 ± .454	1.30 ± .243	.75 ± .116
	Av. gain, retention over final	.67 ± .582	1.02 ± .292	.34 ± .190
	Difference gain	2.69	.08	.41
	P. E. of difference	.739	.380	.223
	Critical Ratio	3.64	.21	1.39
	Part of test	II	III	IV
Method C	Av. gain, final over pretest	2.60 ± .472	1.66 ± .267	.56 ± .093
	Av. gain, retention over final	1.21 ± .596	.66 ± .316	.45 ± .152
	Difference in gain	1.39	1.00	.11
	P. E. of difference	.760	.414	.178
	Critical Ratio	1.8	2.4	.62

TABLE VIII

A TABLE SHOWING THE DIFFERENCES BETWEEN AVERAGE GAINS MADE ON THE PARTS OF THE TEST BY THE STUDY GROUPS, THE PROBABLE ERRORS OF THE DIFFERENCES AND THE CRITICAL RATIOS

		Data from Table II			Data from Table III			Data from Table IV		
Final test over pretest	Part of test	II	III	IV	II	III	IV	II	III	IV
	Av. gain Method A	2.80	1.31	.42	2.80	2.13	1.62	2.80	2.47	1.99
	Av. gain Method B	3.36	1.30	.75	3.36	2.11	2.88	3.36	2.45	3.38
	Difference (A-B)	-.56	.01	-.33	-.56	.02	-1.26	-.56	.02	-1.39
	P. E. of difference	.646	.378	.148	.748	.748	.748	.624	.624	.624
	Critical Ratio	-.85	.029	-2.2	-.75	-.27	-1.68	-.90	.032	-2.2
	Av. gain Method A	2.80	1.31	.42	2.80	2.13	1.62	2.80	2.47	1.99
	Av. gain Method C	2.60	1.66	.56	2.60	2.69	2.16	2.60	3.14	2.52
	Difference (A-C)	.20	-.35	-.14	.20	-.56	-.54	.20	-.67	-.53
	P. E. of difference	.659	.407	.131	.748	.748	.748	.624	.624	.624
	Critical Ratio	.03	-.86	-1.07	.27	-.75	-.72	.33	-1.11	-.86
	Av. gain Method B	3.36	1.30	.75	3.36	2.11	2.88	3.36	2.45	3.38
	Av. gain Method C	2.60	1.66	.56	2.60	2.69	2.16	2.60	3.14	2.52
	Difference (B-C)	.76	-.36	.19	.76	-.58	.72	.76	-.69	.86
	P. E. of difference	.655	.361	.149	.748	.748	.748	.624	.624	.624
	Critical Ratio	.116	-1.0	1.27	1.01	-.77	.96	1.26	-1.15	1.42

		Data from Table II			Data from Table III			Data from Table IV		
Retention test over final test	Part of test	II	III	IV	II	III	IV	II	III	IV
	Av. gain Method A	1.53	.76	.43	1.71	1.23	1.65	1.71	1.43	1.83
	Av. gain Method B	.67	1.02	.34	.67	1.65	1.31	.67	1.93	1.53
	Difference (A-B)	.86	-.26	.09	1.04	-.42	.34	1.04	-.50	.30
	P. E. of difference	.830	.453	.236	.748	.748	.748	.624	.624	.624
	Critical Ratio	1.04	-.57	.38	1.38	-.56	.45	1.7	-.82	.48
	Av. gain Method A	1.53	.76	.43	1.71	1.23	1.65	1.71	1.43	1.83
	Av. gain Method C	1.21	.66	.45	1.21	1.07	1.73	1.21	1.24	2.02
	Difference (A-C)	.32	.10	-.02	.50	.16	-.08	.50	.19	-.19
	P. E. of difference	.480	.469	.207	.748	.748	.748	.624	.624	.624
	Critical Ratio	.38	.21	-.096	.67	.21	-.11	.82	.31	-.31
	Av. gain Method B	.67	1.02	.34	.67	1.65	1.31	.67	1.93	1.53
	Av. gain Method C	1.21	.66	.45	1.21	1.07	1.73	1.21	1.24	2.02
	Difference (B-C)	-.44	.36	-.11	-.54	.58	-.42	-.44	.69	-.49
	P. E. of difference	.833	.430	.243	.748	.748	.748	.624	.624	.624
	Critical Ratio	-.53	.84	-.45	-.72	.77	-.56	-.73	1.15	-.81

COMPARISON OF TWO METHODS OF STUDYING HISTORY

By John J. Dynes
Western State College
Gunnison, Colorado

The problem of this investigation was to study the relative effectiveness of two study techniques in history. There were two different phases to the problem. One phase was to discover if either technique was superior as judged from the amount of material learned in a given length of time. The second phase was to discover if either of the two study techniques was superior as judged from the amount of the learned material that was retained after a given length of time. A minor problem of the investigation was to find how pupils studied and to discover whether or not these study habits influenced in any way the results obtained by the study techniques used in this investigation.

The two study techniques used in this investigation will be designated as Method X and Method Y.

1. Method X:
   Read and reread the material.
2. Method Y:
   (a) Give the material a rapid reading.
   (b) Reread the material, underline essential parts, take notes.
   (c) Review underlined portions and notes.
   (d) Write a brief summary of material read.
   (e) Recall what has been read.

Investigations made of the study habits of pupils show that pupils waste time in school because they do not know how to study. There are many different sources of waste in the process of studying. Time may be wasted by slow movements, by aimless work, by the lack of a definite objective, by not knowing what one is trying to do, or by the use of wrong methods of study. Many methods of study prevent the enjoyment of school, and want of enjoyment of school is likely to result in distaste for intellectual employment in general. McMurry states that; "Bad methods of study easily become a serious factor in adult life, acting as a barrier to one's growth and general usefulness."

A survey of the investigations made on the study habits of pupils reveal a great ignorance on the part of pupils of the whole subject of study. It is evident that children are little acquainted with proper methods of study. These investigations show that pupils have little knowledge of the most economical methods of studying but through the use of trial-and-error methods hope in some way to attain the results desired by the teacher. Teachers themselves have very little definite, scientific knowledge of the relative values of different study techniques. Controlled experiments to determine the relative effectiveness of different methods of studying have not been widely undertaken. Investigations have been made in the different fields of subject matter and on different academic levels, but comparatively few of the methods used in our schools have been tested scientifically to determine their value as teaching techniques. The relative values of different methods used in teaching is not known although we find many different opinions expressed regarding the particular methods that should be used in teaching the different subjects. Many teachers of history contend that the best method of teaching history is that of taking notes while preparing the lesson. Other teachers contend that taking notes is a waste of time and that just as good results are obtained by the method of reading and rereading the assigned material. This study was made in order to try to answer this question with reference to two study techniques in history.

Four different high schools were used in the carrying on of this investigation. The pupils, 134 in number, from two high schools

were used for the purpose of standardizing the test material. The pupils, 144 in number, from the other two high schools were used in the major experiment to test the relative value of the two study techniques.

The material used in this investigation was selected from readings in social science for junior high school pupils. The two different sets of material were of approximately the same length and the content of the same general nature. The words were of about the same difficulty according to the Thorndike word list. A study habit check list was used in order to determine as objectively as possible how the pupils studied. A record was made of all the different activities of each pupil during the period when he was studying. Objective tests were used to determine the amount of material learned and the amount retained. The tests consisted of true-false, multiple-choice and matching questions.

An investigation of how pupils studied was first conducted. Each pupil was asked to study a selected piece of material by his own method of studying. After this was completed the major investigation was started. Each pupil was tested by both methods, X and Y. Half of the pupils first used Method X followed by Method Y, the other half of the pupils used Method Y first followed by Method X. Two days intervened between the use of the two methods. The test for retention was given three weeks later in order to determine the amount of learning by the two methods. A test was given before studying, and again after studying the material. The difference in the scores was used to measure the amount learned during the study period. Thirty minutes were allowed for studying the material, and fifteen minutes for the pretest and final test.

The work was conducted on the individual plan, only one pupil being in the laboratory under the observation of the examiner during the study period. When the pupil entered the study room he was shown to his seat and the material with which he was to work placed on his desk. The examiner in charge of the experiment had an instruction sheet which he read with the pupil. The instruction sheet gave the directions for taking the tests, the time allowed for study, and the method of study to be used during the period.

The two sets of study materials used in the investigation materials were alternated in order to offset any carry-over from one method to the other, and to equalize if not eliminate practice effects. Material A was used half the time with Method X and half the time with Method Y, and Material B was used half the time with Method X and half the time with Method Y.

An analysis of the data collected revealed the following facts: Pupils use the method of reading and rereading the material in preference to any other method; the rate of reading varies widely among high school pupils; pupils follow the technique of study suggested by the teacher when working in the study hall under supervision but when left to themselves will select the method that appeals to them individually. Only 33 pupils out of 190 took notes when allowed to choose their own method of study. (Table I, see following page).[1] Of these 33 pupils, 16 did better work with Method X; 14 did better work with Method Y, and 3 pupils obtained the same results with both methods. This would seem to indicate that the study habits of the pupils had little or no effect upon the results obtained by the different methods used in the major investigation.

The data collected show that of the 144 pupils who participated in the final experiments, 67 learned more with Method X, 73 with Method Y, and 4 pupils made equal gains with both methods. For retention Method X proved to be the better for 49 pupils, and Method Y was better for 76, while 9 pupils retained equal amounts with either method. The average score gained in school G with Method X was 13.42 and the average score gained with Method Y was 14.96. The difference in the mean improvement of Y over X was 1.54 points which is statistically significant. (Table II, see following page). The average score retained with Method X was 9.04 and the average score retained with Method Y was 10.08. The difference in the average amount retained was 1.04 points (Table II) and this amount is also statistically significant. In school H the average score gained with method X was 8.52 (Table III, see following page); and

1. Only three tables chosen from the original data are here presented.

TABLE I

SUMMARY OF THE DATA COLLECTED BY THE USE OF THE "STUDY HABIT CHECK LIST"

Methods of Study	School G (92 pupils)				School H (52 pupils)			
	Yes	No	% Yes	% No	Yes	No	% Yes	% No
1. Pupils went to work at once	92		100		52		100	
2. Rapid survey of work	10	82	11	89		52		100
3. Rapid survey of reading		92		100		52		100
4. Intensive first reading[1]								
5. Times pupil read material[2]								
6. Re-read isolated parts	14	78	15	85		52		100
7. Underline words, phrases	6	86	7	93		52		100
8. List words, dates, events	7	85	8	92	1	51	2	98
9. Take notes on material	7	85	8	92	4	48	8	92
10. Outline material	8	84	9	91	3	49	6	94
11. Written abstract or summary		92		100		52		100
12. Use self devised questions		92		100		52		100
13. Study notes taken (85 took no notes)	5	2	71	29	3	1	75	25
14. Use dictionary, maps, etc.		92		100	9	43	17	82
15. Use "whole method" of study	84	8	91	9	46	6	88	12
16. Use "part method" of study	8	84	9	91	6	46	12	88
17. Use method of recall	9	83	10	90	5	47	10	90
18. Concentrate on work	92		100		52		100	
19. Systematic in study habits	90	2	98	2	52		100	
20. Annoyed by work		92		100		52		100
21. Read and re-read, reg. method used	68	24	74	26	29	23	56	44
22. Ordinarily took notes when studying	24	68	26	74	23	29	44	56
23. Method used: read only	85	7	92	8	48	4	92	8
24. Method used: read and took notes	7	85	8	92	4	48	8	92

Note:--(High School H) [1]Range (6 to 22) 16 minutes. Average: 12.1 minutes.
[2]Twenty-five pupils read the material only once; 27 read the whole or part of the material the second time.

(High School G) [1]Range (6 to 15) 9 minutes. Average: 11.08 minutes.
[2]Fourteen pupils read the material only once; 78 read the whole or part of the material the second time.

TABLE II

A SUMMARY OF THE SCORES MADE BY THE PUPILS USING THE TWO METHODS OF STUDY IN SCHOOL G (92 pupils)

Pretest	Average	S.D.	σ_M	Actual Difference
Method X:	36.72	9.03	.4105	
Method Y:	36.14	9.27	.4214	
Final Test				
Method X:	50.14	6.87	.3123	
Method Y:	51.10	7.71	.3505	
Points Gained				
Method X:	13.42	8.97	.4075	
Method Y:	14.96	10.17	.4634	1.54
	C.R., 4.186			
Retention Scores				
Method X:	45.76	6.88	.3127	
Method Y:	46.22	7.12	.3236	
Points Retained				
Method X:	9.04	9.77	.4438	
Method Y:	10.08	9.82	.4447	1.04
	C.R., 3.048			

TABLE III

A SUMMARY OF THE SCORES MADE BY THE PUPILS USING THE TWO METHODS OF STUDY IN SCHOOL H (52 pupils)

Pretest	Average	S.D.	σ_M	Actual Difference
Method X:	34.75	4.45	.1573	
Method Y:	33.79	3.80	.1343	
Final Test				
Method X:	43.27	5.92	.2093	
Method Y:	42.41	5.73	.2026	
Points Gained				
Method X:	8.52	5.76	.2038	
Method Y:	8.62	5.23	.1493	.10
	C.R., .4746			
Retention Scores				
Method X:	38.79	5.86	.2072	
Method Y:	39.12	5.12	.1810	
Points Retained				
Method X:	4.04	5.55	.1684	
Method Y:	5.33	4.95	.2015	1.29
	C.R., 4.92			

the average score gained with method Y was 8.62. The difference in the mean improvement of Y over X was +0.10. The average retention score for method X was 4.04; the average retention score for method Y was 5.33. The difference in the average amount retained for Y over X was 1.29, and this amount is statistically significant. Three of the four comparisons thus favor method Y.

CONCLUSIONS

The results of this investigation in the light of the data obtained are presented in the form of answers to questions as follows:

A. How do pupils study and what effect do the study habits of pupils have on the results obtained by the study techniques used in this investigation?
   1. Pupils use the whole method of studying.
   2. Pupils use the method of reading and rereading in preference to other methods.
   3. The method of study used in the regular work in history appears to have little, if any effect upon the results obtained from the methods used in this investigation.
   4. The preference for a given method of study appears to have little if any effect upon the results obtained when other methods are employed.

B. What do the results obtained from the investigation of the relative effectiveness of the two different study techniques in history indicate? The data collected in this investigation would seem to justify the following conclusions:
   1. The method of taking notes, outlining the material, summarizing the data, and reviewing, is slightly superior to the method of reading and rereading for immediate learning, both in the total amount learned and the number of pupils who excelled by using this method.
   2. The method of taking notes, outlining the material, summarizing the data, and reviewing the material, is superior to the method of reading and rereading for retention, both in the total amount retained and the number of pupils who excelled by using this method.